Discovery and Characterization of Gut Microbial Enzymes Involved in Xenobiotic Metabolism

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37365731

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Discovery and characterization of gut microbial enzymes involved in xenobiotic metabolism

A dissertation presented by

Vayu Maini Rekdal

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University, Cambridge, Massachusetts

May 2020
Discovery and characterization of gut microbial enzymes involved in xenobiotic metabolism

Abstract

The trillions of microorganisms inhabiting the human gastrointestinal tract (the gut microbiota) harbor metabolic capabilities that expand the range of chemistry taking place in the human body. Gut microbial metabolism directly modifies the chemical structures of diverse ingested and endogenous compounds, making the gut microbiota a key player in pharmacology, nutrition, and host biology. However, few gut microbial transformations have been connected to specific strains and enzymes, limiting our ability to harness microbial chemistry for human health. This thesis describes the discovery and characterization of enzymes that break down the drug L-dopa, catecholamine neurotransmitters, and dietary phytochemicals. This work has provided opportunities to manipulate and study gut microbial metabolism and has uncovered catalytic functions and metabolic pathways that may be broadly relevant to microbial life.

Chapter 2 describes the discovery of organisms and enzymes that metabolize Levodopa (L-dopa), the main treatment for Parkinson’s disease. This drug alleviates many of the symptoms associated with Parkinson’s disease but suffers from high interindividual variability in both efficacy and side effects. Although the gut microbiota has been implicated in L-dopa efficacy and was known to metabolize L-dopa, the microbial enzymes responsible were poorly understood. A major pathway for microbial L-dopa metabolism involves decarboxylation of L-dopa into
dopamine followed by dehydroxylation into \textit{m}-tyramine. We used genome mining, genetics, and biochemistry to identify and characterize a promiscuous tyrosine decarboxylase (TyrDC) from the human gut microbe \textit{Enterococcus faecalis} that directly decarboxylates L-dopa into dopamine. This gene predicted L-dopa metabolism by microbial strains and by complex gut microbiota samples from Parkinson’s patients and neurologically healthy individuals. We next evaluated the activity of known host decarboxylase inhibitors in microbial L-dopa decarboxylation. In contrast to its potency towards host metabolism, the commonly prescribed decarboxylase inhibitor carbidopa only weakly microbial L-dopa decarboxylation, suggesting that some metabolism may be unaccounted for in the current treatment regimen. This observation led us to identify (S)-α-fluoromethyltyrosine, a small molecule that selectively inhibits microbial L-dopa decarboxylation by isolated strains and microbial communities. Finally, we uncover that the gut bacterium \textit{Clostridium sporogenes} uses enzymes involved in aromatic amino acid metabolism to degrade L-dopa via an alternative pathway.

Chapter 3 describes our work identifying and characterizing microbial strains, genes, and enzymes involved in dehydroxylating the catechol ring of dopamine, a human neurotransmitter and the intermediate in the two-step microbial metabolism of L-dopa into \textit{m}-tyramine. We used enrichment culturing from complex human fecal samples to isolate a strain of \textit{Eggerthella lenta} that dehydroxylates dopamine to give \textit{m}-tyramine. We then used RNA-sequencing, comparative genomics, chemical genetics, and in vitro biochemistry to identify a molybdenum-dependent enzyme (named the dopamine dehydroxylase, or Dadh) involved in this reaction. Bioinformatics analysis revealed that a single nucleotide polymorphism in the gene encoding this enzyme correlated with metabolism by isolated strains and complex microbial communities.
Chapter 4 expands on our discovery of Dadh. Dopamine dehydroxylation is just one of many examples of catechol dehydroxylation, a poorly understood but prominent transformation in gut microbial metabolism of endogenous compounds, dietary small molecules, and pharmaceutical drugs. We first characterized the substrate scope and regulation of Dadh and found a remarkable specificity for the catecholamine scaffold. We also demonstrated that dopamine can serve as an alternative electron acceptor for *E. lenta*, providing a potential physiological explanation for this high specificity. To identify strains capable of metabolizing catechols beyond dopamine, we screened a collection of gut bacteria for metabolism of catecholic drugs, dietary compounds, and siderophores. This screen unveiled that *Gordonibacter pamelaeae* 3C dehydroxylates DOPAC and that the closely related *E. lenta* dehydroxylated diet-derived hydrocaffeic acid and (+)-catechin. RNA-sequencing and comparative genomics revealed candidate molybdenum-dependent enzymes involved in these reactions, and the native purification of the hydrocaffeic acid dehydroxylase (Hcdh) strongly supported a model in which each catechol is metabolized its own dedicated gut microbial enzyme. Phylogenetic and sequence similarity network analyses established catechol dehydroxylases as a new class of molybdenum-dependent enzymes harboring vast uncharacterized diversity among gut microbes and environmental microbes alike. Finally, we found that catechol dehydroxylation was present among gut microbiotas of mammals representing diverse diets and phylogenetic origins, reinforcing that this activity can take place in habitats beyond the human gut. This suggested that the chemical strategies used to enable microbial survival and interactions in the human gut are relevant to a broad range of species and habitats.

Chapter 5 describes our progress towards mechanistic and biochemical characterization of catechol dehydroxylases. Based on the Dadh and Hcdh substrate scope and phylogenetic analysis in Chapter 4, we proposed a mechanism for catechol dehydroxylation that represented a new
strategy for aromatic dehydroxylation. Deuterium incorporation experiments investigating hydrocaffeic acid dehydroxylation in cell lysates supported our mechanistic proposal, revealing that a single deuterium was incorporated at high levels (>95%) into the aromatic ring of the dehydroxylated product. To enable further mechanistic and biochemical characterization, we turned to heterologous expression. We cloned and expressed the gut microbial dopamine dehydroxylase (Dadh), hydrocaffeic acid dehydroxylase (Hcdh), catechin dehydroxylase (Cadh), DOPAC dehydroxylase (Dodh), catechol lignan dehydroxylase (Cldh), and a putative dehydroxylase from the soil microbe T. aromatica (Tardh). Our efforts involved extensive testing of different plasmids across different heterologous and media conditions. However, no constructs yielded soluble, active protein, suggesting further work is necessary to overcome the challenges associated with the complex maturation and assembly of molybdenum-dependent enzymes.
# Table of Contents

Abstract ................................................................................................................................. iii  
Table of Contents .................................................................................................................. vii  
List of Figures ....................................................................................................................... xxii  
List of Tables ......................................................................................................................... xxxix  
List of Abbreviations ........................................................................................................... xlii  
Acknowledgements .............................................................................................................. xlviii  

## Chapter 1. Gut microbial metabolism of xenobiotics and endogenous compounds .... 1  
1.1 The human microbiota ................................................................................................. 2  
1.2 The human gut microbiota and its involvement in metabolism .................................. 2  
1.3 Microbial and host metabolism of xenobiotics ............................................................ 6  
1.4 Contrasting the chemistry involved in host and microbial xenobiotic processing ....... 7  
  1.4.1 Host enzymes involved in xenobiotic metabolism ....................................................... 8  
  1.4.2 Hydrolytic reactions ................................................................................................... 9  
  1.4.3 Lyases ....................................................................................................................... 10  
  1.4.4 Reductive transformations ......................................................................................... 13  
  1.4.5 Functional group transfer reactions .......................................................................... 15  
  1.4.6 Radical chemistry ..................................................................................................... 16  
  1.4.7 Uncharacterized microbial chemistry ...................................................................... 19  
1.5 Metabolism of dietary compounds ............................................................................. 20  
  1.5.1 Dietary protein ......................................................................................................... 21  
  1.5.2 Dietary lipids ........................................................................................................... 22  
  1.5.3 Dietary phytochemicals ........................................................................................... 24
Chapter 2. Discovery, characterization, and inhibition of gut microbial enzymes involved in the metabolism of L-dopa, the main treatment for Parkinson’s disease ................. 115

2.1 Introduction.................................................................................................................. 116

2.2 Results.......................................................................................................................... 124

2.2.1 Discovery of strains that decarboxylate L-dopa ......................................................... 124

2.2.2 The human gut bacterium Enterococcus faecalis decarboxylates L-dopa using a promiscuous PLP-dependent tyrosine decarboxylase ................................................. 131

2.2.3 Influence of pH and tyrosine concentration on gut microbial L-dopa decarboxylation ......................................................................................................................... 135

2.2.4 L-dopa decarboxylation promotes survival of E. faecalis in acidic conditions ............................................. 140

2.2.5 TyrDC abundance correlates with L-dopa metabolism ex vivo ..................................... 141

2.2.6 Clostridium sporogenes metabolizes L-dopa using enzymes involved in aromatic amino acid fermentation ........................................................................................................... 148

2.2.7 Evaluation of human aromatic amino acid decarboxylase inhibitors towards gut microbial L-dopa decarboxylation ..................................................................................... 152

2.2.8 The small-molecule inhibitor (S)-α-fluoromethyltyrosine selectively targets bacterial decarboxylation .............................................................................................................. 157

2.3 Discussion ...................................................................................................................... 166

2.4 Methods ......................................................................................................................... 169

2.4.1 General materials ....................................................................................................... 169

2.4.2 General methods ......................................................................................................... 169
2.4.3 LC-MS methods ........................................................................................................ 170
2.4.4 BLASTP search for \textit{L. brevis} TyrDC in Human Microbiome Project reference isolate genomes ................................................................................................................................. 173
2.4.5 Assessment of conservation of TyrDC across \textit{E. faecalis} genomes ...................... 173
2.4.6 Screen for L-dopa decarboxylation in anaerobic bacterial cultures....................... 174
2.4.7 Confirmation of \textit{E. faecalis} MMH594 tyrDC mutants ............................................. 175
2.4.8 Cloning and expression of \textit{E. faecalis} MMH594 tyrosine decarboxylase ............ 175
2.4.9 Michaelis–Menten kinetics of \textit{E. faecalis} TyrDC ....................................................... 176
2.4.10 Competition of L-dopa and tyrosine for reaction with purified \textit{E. faecalis} TyrDC ........................................................................................................................................ 177
2.4.11 Assessment of the impact of pH on L-dopa and tyrosine decarboxylation by \textit{E. faecalis} ........................................................................................................................................ 177
2.4.12 Assessment of the impact of pH and tyrosine on L-dopa decarboxylation by \textit{E. faecalis} MMH594 ........................................................................................................................................ 178
2.4.13 Assessment of the impact of pH and tyrosine on L-dopa decarboxylation by three fecal samples from neurologically healthy humans ......................................................................................................................... 178
2.4.14 Calculation of drug concentrations in the small intestine .................................... 179
2.4.15 Cloning and heterologous expression of \textit{H. sapiens} aromatic amino acid decarboxylase ........................................................................................................................................ 179
2.4.16 Evaluation of inhibitors towards \textit{E. faecalis} and \textit{H. sapiens} decarboxylases ..... 180
2.4.17 Identification of a PLP-inhibitor adduct upon incubation of enzymes with AFMT ................................................................................................................................. 182
2.4.18 Substrate screen with \textit{E. faecalis} TyrDC and \textit{H. sapiens} AADC ....................... 182
2.4.19 Evaluation of inhibitors towards growth and L-dopa decarboxylation by E. faecalis ................................................................. 182

2.4.20 Collection of fecal samples from neurologically healthy human patients........ 183

2.4.21 Collection of fecal samples from Parkinson’s disease patients.................... 183

2.4.22 Ex vivo assays with d3-phenyl-L-dopa .................................................. 184

2.4.23 qPCR assays ......................................................................................... 184

2.4.24 Metagenomic analysis of E. faecalis/tyrDC abundance and prevalence across human patients .......................................................................................................................... 185

2.4.25 MTT assay for HeLa cell viability in the presence of AFMT ......................... 185

2.4.26 Pharmacokinetic experiment in gnotobiotic mice......................................... 185

2.4.27 Extraction and analysis of L-dopa from tail vein blood ............................... 186

2.4.28 Incubation of wild-type and mutant C. sporogenes strains with d3-phenyl-L-dopa. ................................................................................................................................. 187

2.5 References ......................................................................................................... 188

Chapter 3. Discovery of gut microbial strains and enzymes that dehydroxylate dopamine .......................................................................................................................... 199

3.1 Introduction ........................................................................................................ 200

3.2 Results ................................................................................................................ 207

3.2.1 Screen for dopamine dehydroxylation among human gut microbes ............. 207

3.2.2 Enrichment culturing from a human fecal sample identifies the dopamine dehydroxylating organism Eggerthella lenta A2 .......................................................... 209

3.2.3 Dopamine dehydroxylation is inducible in Eggerthella lenta A2 .................... 212

3.2.4 Optimization of culture conditions for RNA-sequencing in Eggerthella lenta A2 215
3.2.5 RNA-sequencing identifies a putative molybdenum-dependent dopamine dehydroxylase ................................................................. 218

3.2.6 Native purification of the dopamine dehydroxylase from *E. lenta* A2 ............... 225

3.2.7 A single nucleotide polymorphism (SNP) discriminates dopamine metabolizing *E. lenta* strains from non-metabolizers ............................................................. 232

3.2.8 The dopamine dehydroxylase SNP status predicts dopamine dehydroxylation by complex gut microbiota communities ........................................................................ 240

3.2.9 *E. faecalis* and *E. lenta* produce *m*-tyramine from L-dopa in co-culture and in complex gut microbiota communities ......................................................... 245

3.2.10 Carbidopa and AFMT do not impact *E. lenta* growth or dopamine dehydroxylation ......................................................................................... 249

3.2.11 Attempts to reconstitute gut bacterial metabolism of L-dopa to *m*-tyramine in gnotobiotic mice ......................................................................................... 251

3.3 Discussion .................................................................................................................................................................................. 262

3.4 Methods ..................................................................................................................................................................................... 266

3.4.1 General materials ................................................................................................................................................................. 266

3.4.2 General methods ................................................................................................................................................................. 266

3.4.3 LC-MS methods ................................................................................................................................................................. 267

3.4.4 Screen of a collection of gut bacteria in anaerobic culture for dopamine dehydroxylation using a colorimetric assay ......................................................... 269

3.4.5 Collection of fecal samples from neurologically healthy human patients ........... 269

3.4.6 Collection of fecal samples from Parkinson’s disease patients ............................ 270
3.4.7 Enrichment culturing and isolation of a dopamine dehydroxylating gut bacterial strain
............................................................................................................................................. 270
3.4.8 Culture-based assays to assess the inducibility of dopamine dehydroxylation in *E. lenta* A2............................................................................................................................................. 273
3.4.9 Lysate assays to assess dopamine dehydroxylation in *E. lenta* A2 ...................... 273
3.4.10 Evaluation of the impact of arginine and formate on growth and metabolism by *E. lenta* .......................................................................................................................................................................................... 274
3.4.11 Screen of gut Actinobacterial isolates for dopamine dehydroxylation............. 275
3.4.12 RNA-sequencing in *E. lenta* A2, *E. lenta* DSM2243, and *E. lenta* 28b ......... 275
3.4.13 Evaluation of the effects of tungstate and molybdate on dopamine dehydroxylation in cultures of *E. lenta* A2 ............................................................................................................................................. 277
3.4.14 Distribution of *E. lenta* A2 dadh among sequenced genomes....................... 277
3.4.15 Assessment of the conservation of *dadh* and surrounding genes across a gut Actinobacterial library .......................................................................................................................................................................................... 277
3.4.16 Conservation of *Eggerthella lenta* A2 dadh among Actinobacterial isolates and alignment of protein sequences.......................................................................................................................................................................................... 277
3.4.17 Ultracentrifugation of *E. lenta* A2 lysates to assess localization of dopamine dehydroxylase activity .......................................................................................................................................................................................... 278
3.4.18 Precipitation of *E. lenta* A2 lysates using ammonium sulfate ....................... 279
3.4.19 Anaerobic activity-based purification of *E. lenta* A2 Dadh ............................... 280
3.4.20 Co-culturing of *E. faecalis* and *E. lenta* A2 .............................................. 284
3.4.21 Gain-of-function assays in complex human gut microbiota samples ............ 285
3.4.22 qPCR assays on DNA from bacterial communities....................................... 285

xiii
3.4.23 Ex vivo assays for dopamine dehydroxylation by LC-MS and Sanger sequencing .............................................................. 286
3.4.24 Metagenomic analysis of *E. lenta/dadh* and *E. faecalis/tyrDC* abundance and prevalence across human patients ......................................................... 287
3.4.25 Evaluation of inhibitors towards growth and dopamine metabolism in *E. lenta*. 287
3.4.26 Gnotobiotic mouse experiment with *E. faecalis* and *E. lenta* .................. 288
3.5 References .................................................................................................................. 295

Chapter 4. A widely distributed metalloenzyme class enables gut microbial metabolism of host-and diet-derived catechols ................................................................. 307

4.1 Introduction ............................................................................................................. 308
4.2 Results .................................................................................................................. 314
  4.2.1 Dadh specifically metabolizes catecholamines that are available in the gut .... 314
  4.2.2 Dopamine promotes gut bacterial growth by serving as an alternative electron acceptor ............................................................... 323
  4.2.3 A screen of human gut Actinobacteria uncovers dehydroxylation of host and plant-derived catechols .............................................................. 335
  4.2.4 Impact of DOPAC, catechin, and hydrocaffeic acid on gut Actinobacterial growth in basal medium .......................................................... 346
  4.2.5 Catechol dehydroxylation by gut Actinobacteria is inducible and specific .... 349
  4.2.6 Discovery of enzymes that dehydroxylate (+)-catechin and hydrocaffeic acid..... 353
  4.2.7 Discovery of a candidate DOPAC dehydroxylase ........................................... 369
  4.2.8 *G. pamelaeae* likely harbors a distinct hydrocaffeic acid dehydroxylase ........ 380
  4.2.9 Norepinephrine and caffeic acid do not induce unique dehydroxylases ....... 381
4.2.10 Model for *Eggerthella* and *Gordonibacter* dehydroxylases .................................................. 385
4.2.11 Catechol dehydroxylases are variably distributed in metagenomes and correlate with metabolism by complex gut microbiota samples ex vivo ................................................................. 389
4.2.12 Catechol dehydroxylases are distinct from other molybdenum-dependent enzymes ................................................................. 394
4.2.13 GNN analysis of *Eggerthella* and *Gordonibacter* dehydroxylases ................................. 403
4.2.14 Catechol dehydroxylases are widely distributed among sequenced microbes .... 408
4.2.15 Catechol dehydroxylase reactivity is present across the gut microbiotas of mammals representing distinct diets and phylogenetic origins ................................................................. 419
4.3 Discussion .............................................................................................................................................. 423
4.4 Experimental methods .......................................................................................................................... 429
4.4.1 General materials and methods ...................................................................................................... 429
4.4.2 LC-MS methods ............................................................................................................................. 431
4.4.3 Colorimetric assay for catechol detection ....................................................................................... 437
4.4.4 Anaerobic activity-based purification of *E. lenta* A2 dopamine dehydroxylase ... 438
4.4.5 Assays of the *E. lenta* A2 dopamine dehydroxylase substrate scope ................................. 440
4.4.6 Metabolism of dopamine analogs by *E. lenta* A2 cells .............................................................. 441
4.4.7 RNA-sequencing experiments with *E. lenta* .............................................................................. 442
4.4.8 Time course of DOPAC metabolism by *G. pamelaeae* 3C .................................................... 443
4.4.9 RNA-sequencing experiments with *G. pamelaeae* 3C ........................................................... 444
4.4.10 Growth of *E. lenta* A2 in BHI medium with and without dopamine ......................... 444
4.4.11 Preparation of basal medium lacking electron acceptors ....................................................... 445
4.4.12 Growth of single *E. lenta* strains in basal medium .............................................................. 445
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.13 Growth of <em>G. pamelaeeae</em> 3C strains in basal medium</td>
<td>446</td>
</tr>
<tr>
<td>4.4.14 Competition of <em>E. lenta</em> strains in basal medium</td>
<td>446</td>
</tr>
<tr>
<td>4.4.15 Human fecal samples used in this study</td>
<td>447</td>
</tr>
<tr>
<td>4.4.16 Growth of human fecal samples in basal medium with dopamine</td>
<td>448</td>
</tr>
<tr>
<td>4.4.17 qPCR assays for <em>E. lenta</em> and <em>dadh</em> abundance in human fecal samples grown in basal medium with and without dopamine</td>
<td>448</td>
</tr>
<tr>
<td>4.4.18 Amplification of full-length <em>dadh</em> and sequencing of the SNP at position 506 from human fecal samples grown in basal medium with and without dopamine</td>
<td>449</td>
</tr>
<tr>
<td>4.4.19 Screen of gut Actinobacteria for metabolism of catechols</td>
<td>449</td>
</tr>
<tr>
<td>4.4.20 Confirmation of dehydroxylation of (+)-catechin and hydrocaffeic acid by <em>E. lenta</em> A2</td>
<td>451</td>
</tr>
<tr>
<td>4.4.21 Confirmation of dehydroxylation of DOPAC by <em>Gordonibacter pamelaeeae</em> 3C</td>
<td>452</td>
</tr>
<tr>
<td>4.4.22 PBS resuspension assays for inducibility and oxygen sensitivity of catechol dehydroxylases</td>
<td>452</td>
</tr>
<tr>
<td>4.4.23 Effect of tungstate on growth and catechol dehydroxylation by <em>E. lenta</em> A2 and <em>G. pamelaeeae</em> 3C</td>
<td>454</td>
</tr>
<tr>
<td>4.4.24 Lysate assays for transcriptional and biochemical specificity of dehydroxylases from <em>G. pamelaeeae</em> 3C and <em>E. lenta</em> A2</td>
<td>455</td>
</tr>
<tr>
<td>4.4.25 Comparative genomics among human gut Actinobacteria</td>
<td>457</td>
</tr>
<tr>
<td>4.4.26 Anaerobic activity-based purification of <em>E. lenta</em> A2 hydrocaffeic acid dehydroxylase</td>
<td>457</td>
</tr>
<tr>
<td>4.4.27 Proteomics of <em>E. lenta</em> A2 hydrocaffeic acid dehydroxylase</td>
<td>459</td>
</tr>
</tbody>
</table>
4.4.28 Assays of partially purified *E. lenta* A2 hydrocaffeic acid dehydroxylase substrate scope ........................................................................................................ 461
4.4.29 Metagenomic analysis of catechol dehydroxylase abundance and prevalence across human patients .................................................................................................................. 462
4.4.30 Incubation of human fecal samples with *d*<sub>L</sub>-dopamine, hydrocaffeic acid, (+)-catechin, and *d*<sub>S</sub>-DOPAC ......................................................................................................................... 463
4.4.31 qPCR assays for *dodh* abundance in human fecal samples grown in BHI medium with *d*<sub>S</sub>-DOPAC ......................................................................................................................... 463
4.4.32 Phylogenetic analysis of relationship between catechol dehydroxylases and other characterized members of the bis-MGD enzyme family ......................................................... 464
4.4.33 Construction of a protein sequence similarity network of the bis-MGD enzyme family. ................................................................................................................................. 465
4.4.34 Gene neighborhood network analysis of *E. lenta* and *Gordonibacter*-type dehydroxylases ....................................................................................................................... 465
4.4.35 Phylogenetic analysis of catechol dehydroxylases encoded by gut Actinobacteria and environmental isolates .............................................................................................. 465
4.4.36 Phylogenetic analysis of relationship between representative dehydroxylase homologs and other characterized members of the bis-MGD enzyme family .......... 467
4.4.37 Mammalian fecal samples used in this study ......................................................... 468
4.4.38 Screen for catechol dehydroxylation by mammalian gut microbiota samples .... 468
4.4.39 Construction of a mammalian phylogenetic tree ................................................... 469
4.5 Materials and methods for synthetic chemistry ....................................................... 470
4.5.1 General materials and methods ............................................................................ 470
4.5.2 Preparation of 2-aminoethylbenzendiol/-triol derivatives ........................................ 472
4.5.3 Preparation of 2-amino-4-(2-aminoethyl)phenol ...................................................... 491
4.5.4 Preparation of hydroxytyrosol ..................................................................................... 495
4.5.5 Additional information .............................................................................................. 496
4.5.6 $^1$H / $^{13}$C SPECTRA .................................................................................................. 497

4.6 References ........................................................................................................................ 539

Chapter 5. Progress towards biochemical and mechanistic characterization of catechol
dehydroxylases ....................................................................................................................... 552

5.1 Introduction ....................................................................................................................... 553

5.2 Results .............................................................................................................................. 585

5.2.1 Lysate assays reveal that hydrocaffeic acid dehydroxylation incorporates high levels
of a single deuterium into the product $m$-hppa ..................................................................... 585

5.2.2 General approach for heterologous expression ............................................................. 589

5.2.3 Expression of Dadh in E. coli ...................................................................................... 589

5.2.4 Expression of Dadh in R. erythropolis and E. faecalis .................................................. 607

5.2.5 Expression of additional catechol dehydroxylases in E. coli and Shewanella sp. ANA-3 .......................................................................................................................... 609

5.3 Discussion .......................................................................................................................... 614

5.4 Methods ............................................................................................................................ 620

5.4.1 General materials ........................................................................................................ 620

5.4.2 General methods .......................................................................................................... 621

5.4.3 LC-MS methods .......................................................................................................... 623

5.4.4 Lysate assays in deuterated buffer in E. lenta A2 ......................................................... 625
| 5.4.5 Cloning of pET28a-\textit{dadhA}_C-His\textsubscript{6} | ........................................................................ | 626 |
| 5.4.6 Cloning of pET28a-\textit{dadhA\_noTATsignal}_C-His\textsubscript{6} | ........................................................................ | 626 |
| 5.4.7 Cloning of pET28a-\textit{dadhA\_noTATsignal}_N-His\textsubscript{6} | ........................................................................ | 626 |
| 5.4.8 Cloning of pET28a-\textit{dadhA\_narG-TAT}_C-His\textsubscript{6} | ........................................................................ | 627 |
| 5.4.9 Cloning of pETDuet1\_\textit{dadhE}_N-His\textsubscript{6} | ........................................................................ | 627 |
| 5.4.10 Cloning of pCOLA-\textit{dadhA}_C-strep\_\textit{dadhB}_N-His\textsubscript{6} | ........................................................................ | 627 |
| 5.4.11 Cloning of pETDuet1\_\textit{dadhC}_N-His\textsubscript{6}\_\textit{dadhE}_C-strep | ........................................................................ | 628 |
| 5.4.12 Cloning of pETDuet1\_\textit{dadhABC}_C-strep\_\textit{dadhE}_N-His\textsubscript{6} | ........................................................................ | 629 |
| 5.4.13 Cloning of pTrcHis2A-\textit{dadhABC}_C-His\textsubscript{6} | ........................................................................ | 630 |
| 5.4.14 Cloning of pTrcHis2A-\textit{dadhAB}_C-His\textsubscript{6} | ........................................................................ | 630 |
| 5.4.15 Cloning of pEXT22-\textit{dadhE} | ........................................................................ | 630 |
| 5.4.16 Cloning of pBAD18K-\textit{dadhDEFGH} | ........................................................................ | 631 |
| 5.4.17 Cloning of pTipQC-2-\textit{dadhABC} | ........................................................................ | 631 |
| 5.4.18 Cloning of pTRKH2-\textit{dadhABC} | ........................................................................ | 631 |
| 5.4.19 Cloning of pBAD18K-\textit{dadhABC} | ........................................................................ | 632 |
| 5.4.20 Cloning of pBAD18K-\textit{hcdhABC} | ........................................................................ | 632 |
| 5.4.21 Cloning of pBAD18K-\textit{dodhAB} | ........................................................................ | 632 |
| 5.4.22 Cloning of pBAD18K-\textit{cldhAB} | ........................................................................ | 633 |
| 5.4.23 Cloning of pBAD18K-\textit{tardh} | ........................................................................ | 633 |
| 5.4.24 Cloning of pTrcHis2A-\textit{dodhAB} | ........................................................................ | 633 |
| 5.4.25 Cloning of pTrcHis2A-\textit{cldhAB} | ........................................................................ | 634 |
| 5.4.26 Cloning of pET28a-\textit{dodhAB}_C-His\textsubscript{6} | ........................................................................ | 634 |
| 5.4.27 Cloning of pETDuet1-\textit{dadhA}_C-His\textsubscript{6} | ........................................................................ | 635 |
5.4.28 Cloning of pETDuet1-dodhA_C-His6_dodhB .................................................. 635
5.4.29 Cloning of pET28a-narG-dodhAB_C-His6 .............................................. 636
5.4.30 Cloning of pETDuet1-narG_dodhA_C-His6_dodhB .................................. 636
5.4.31 Protein expression in E. coli ......................................................................... 636
5.4.32 Purification of DadhA from E. coli BL21(DE3) ........................................... 639
5.4.33 Expression of catechol dehydroxylase homologs in E. coli and Shewanella sp. ANA-3 ........................................................................................................ 640
5.4.34 Protein expression in E. faecalis OG1RF ...................................................... 641
5.4.35 Protein expression in R. erythropolis L-88 .................................................... 641
5.4.36 Lysate assays to assess catechol dehydroxylase activity ............................ 642
5.4.37 LC-MS analysis of culture supernatants ....................................................... 643
5.5 References ......................................................................................................... 644

Chapter 6. Concluding remarks and future directions ............................................. 656

6.1 Concluding remarks .......................................................................................... 657
6.2 Future directions .................................................................................................. 667
  6.2.1 Evaluate microbial contributions to L-dopa efficacy in vivo ....................... 667
  6.2.2 Further characterize AFMT and develop novel analogs ............................. 668
  6.2.3 Understand how catechol dehydroxylation impacts host and gut microbiota in vivo ........................................................................................................ 669
  6.2.4 Understand the molecular basis of catechol dehydroxylase regulation in human gut microbes ................................................................................ 670
  6.2.5 Characterize catechol dehydroxylase diversity in the human gut and beyond ..... 671
  6.2.6 Characterize the mechanism of catechol dehydroxylation .......................... 675
6.3 References........................................................................................................................................ 676
List of Figures

Figure 1.1 Map of the human gastrointestinal tract and locations of gut microbes....................... 3
Figure 1.2 Overview of gut microbial metabolism........................................................................... 5
Figure 1.3 Overview of host xenobiotic metabolism....................................................................... 8
Figure 1.4 Examples of gut microbial hydrolytic reactions............................................................... 10
Figure 1.5 Examples of gut microbial lyase reactions...................................................................... 12
Figure 1.6 Examples of gut microbial reductive transformations...................................................... 14
Figure 1.7 Examples of gut microbial functional group transfer reactions....................................... 16
Figure 1.8 Examples of gut microbial radical chemistry................................................................. 18
Figure 1.9 Examples of uncharacterized gut microbial metabolism.................................................. 20
Figure 1.10 Gut microbial metabolism of amino acids derived from dietary protein. ................... 22
Figure 1.11 Gut microbial metabolism of dietary lipids..................................................................... 23
Figure 1.12 Gut microbial metabolism of dietary phytochemicals...................................................... 27
Figure 1.13 Gut microbial metabolism of molecules produced during cooking............................... 29
Figure 1.14 Gut microbial metabolism of food additives................................................................. 31
Figure 1.15 Gut microbial metabolism of anti-inflammatory and GI agents..................................... 33
Figure 1.16 Gut microbial metabolism of immunosuppressants...................................................... 35
Figure 1.17 Gut microbial metabolism of cancer chemotherapy drugs.......................................... 38
Figure 1.18 Gut microbial metabolism of cardiovascular drugs....................................................... 40
Figure 1.19 Gut microbial metabolism of anti-viral drugs............................................................... 42
Figure 1.20 Gut microbial metabolism of Central Nervous System (CNS) drugs............................ 44
Figure 1.21 Gut microbial metabolism of herbal supplements........................................................ 46
Figure 1.22 Gut microbial metabolism of mucus components........................................................ 49
Figure 1.23 Gut microbial metabolism of bile salts and bile acids................................. 51
Figure 1.24 Gut microbial metabolism of hormones....................................................... 52
Figure 1.25 Gut microbial metabolism of neurotransmitters.......................................... 55
Figure 1.26 Use of high-throughput screens to discover new microbial transformations.......... 61
Figure 1.27 Use of computationally guided approaches to discover new microbial transformations. ............................................................................................................................. 64
Figure 1.28 Use of genetic screens to link transformations to genes and enzymes............. 66
Figure 1.29 Use of chemically guided genome mining to link transformations to genes and enzymes.......................................................................................................................... 68
Figure 1.30 Use of transcriptomics and comparative genomics to link transformations to genes and enzymes.......................................................................................................................... 70
Figure 1.31 Use of native purification to link transformations to genes and enzymes.......... 73
Figure 1.32 Use of high-throughput screening to identify microbiota-targeted inhibitors......... 77
Figure 1.33 Use of rational design to identify gut microbiota-targeted small-molecule inhibitors. .......................................................................................................................... 79
Figure 1.34 FDA approved drugs can inhibit gut microbial enzymes................................. 82
Figure 2.1 Mechanism of action and host metabolism of L-dopa....................................... 118
Figure 2.2 Gut microbial metabolism of L-dopa................................................................. 122
Figure 2.3 General decarboxylation mechanism of PLP-dependent decarboxylases............ 125
Figure 2.4 Phylogenetic distribution of tyrosine decarboxylases (TyrDC) in the human microbiota. .......................................................................................................................... 126
Figure 2.5 Phylogenetic tree based on the alignment of the TyrDC protein sequences from Enterococcal HMP reference genomes................................................................. 127
Figure 2.6 Testing representative gut microbial strains encoding TyrDC reveals that *E. faecalis* strains reproducibly convert L-dopa to dopamine. ................................................................. 127

Figure 2.7 Presence of tyrosine decarboxylase homologs across Enterococci and related strains. ..................................................................................................................... 128

Figure 2.8 L-dopa decarboxylation activity among Enterococci and related strains.............. 130

Figure 2.9 Genomic organization of the tyrDC operon in *E. faecalis* strains used in this study. 132

Figure 2.10 *E. faecalis* MMH594 tyrDC mutants do not decarboxylate L-dopa. ................. 133

Figure 2.11 Anaerobic growth of wild-type (WT) *E. faecalis* MMH594 and a tyrDC mutant. 134

Figure 2.12 The *E. faecalis* tyrosine decarboxylase decarboxylates L-dopa and tyrosine. .... 134

Figure 2.13 L-dopa and tyrosine competition experiment with TyrDC. .................................. 136

Figure 2.14 Anaerobic growth, L-dopa metabolism, and tyrosine metabolism of four *E. faecalis* strains across varying pH. ......................................................................................... 137

Figure 2.15 Impact of pH and tyrosine concentration on L-dopa metabolism by *E. faecalis* MMH594 grown in BHI medium. .............................................................. 138

Figure 2.16 Impact of pH and tyrosine concentration on d3-phenyl-L-dopa metabolism by three human fecal samples grown in BHI medium......................................................... 139

Figure 2.17 L-dopa decarboxylation promotes acid survival of *E. faecalis*. ......................... 141

Figure 2.18 Metabolism of d3-phenyl-L-dopa by gut microbiota samples from 19 unrelated neurologically healthy individuals ex vivo. .............................................................. 143

Figure 2.19 tyrDC and *E. faecalis* abundance correlates with L-dopa metabolism status. .... 143

Figure 2.20 Correlation between tyrDC and *E. faecalis* abundance in ex vivo cultures incubated with L-dopa.................................................................................................. 144
Figure 2.21 Correlation between tyrDC and E. faecalis abundance in human gut metagenomes. ........................................................................................................................................................................... 144

Figure 2.22 Detection of alternative L-dopa metabolites in ex vivo cultures from 12 unrelated neurologically healthy human subjects........................................................................................................................................................................................................... 145

Figure 2.23 Metabolism of d3-phenyl-L-dopa by stool samples from 12 unrelated Parkinson’s disease patients......................................................................................................................................................................................................................... 147

Figure 2.24 tyrDC and E. faecalis abundance correlates with L-dopa metabolism status in Parkinson’s patients........................................................................................................................................................................................................................................... 147

Figure 2.25 Pathway for tyrosine metabolism by C. sporogenes and proposed pathway for alternative L-dopa metabolism. ........................................................................................................................................................................................................................................... 149

Figure 2.26 C. sporogenes metabolizes L-dopa to hydrocaffeic acid............................................. 151

Figure 2.27 Human AADC inhibitors and their mechanism of action............................................. 153

Figure 2.28 Effect of benserazide, benserazide free hydrazine, and carbidopa on L-dopa decarboxylation by AADC, TyrDC and E. faecalis. ................................................................................................................................................................................................................................. 154

Figure 2.29 Carbidopa does not inhibit growth of E. faecalis............................................................... 155

Figure 2.30 Time course of degradation of d3-phenyl-L-dopa in the presence and absence of 2 mM carbidopa in a stool sample from a neurologically healthy patient. .............................................................................................................. 156

Figure 2.31 Activity of carbidopa in human fecal microbiota samples................................................. 156

Figure 2.32 Examples of warheads used to inhibit PLP-dependent decarboxylases and mechanism of action of α-fluoromethyl warheads. ........................................................................................................................................................................................................................................... 158

Figure 2.33 Screen of amino acid substrates for AADC and TyrDC. ..................................................... 160

Figure 2.34 Effect of AFMT on L-dopa decarboxylation by AADC, TyrDC and E. faecalis. 161

Figure 2.35 Covalent inhibition of E. faecalis TyrDC by AFMT. .............................................................. 162
Figure 2.36 AFMT does not inhibit growth of *E. faecalis* ........................................ 163
Figure 2.37 Activity of AFMT in human fecal microbiota samples .............................. 163
Figure 2.38 MTS assay evaluating AMFT toxicity toward HeLa cells ........................... 164
Figure 2.39 Effect of AFMT on serum L-dopa in gnotobiotic mice colonized with *E. faecalis* MMH594 and given L-dopa and carbidopa ............................................................... 165
Figure 3.1 Metabolism and physiological effects of dopamine ...................................... 205
Figure 3.2 Reaction performed by 4-hydroxybenzoyl CoA reductase ............................ 206
Figure 3.3 The Arnow method for colorimetric detection of dopamine in bacterial cultures ... 208
Figure 3.4 Enrichment culturing for isolating dopamine dehydroxylating strains from complex human stool samples .......................................................... 210
Figure 3.5 *Eggerthella lenta* A2 quantitatively and regiospecifically dehydroxylates dopamine to *m*-tyramine in anaerobic culture ......................................................... 211
Figure 3.6 Culture experiments establishes that dopamine dehydroxylation is inducible in *Eggerthella lenta* A2 .............................................................. 213
Figure 3.7 Assays with *E. lenta* A2 lysates suggest dopamine dehydroxylation is an inducible and oxygen sensitive enzymatic activity .............................................. 214
Figure 3.8 Impact of arginine on growth and dopamine dehydroxylation in *E. lenta* A2 when dopamine is added at the time of inoculation ........................................ 216
Figure 3.9 Impact of arginine and formate dopamine dehydroxylation *E. lenta* A2 when dopamine is added during exponential growth .............................................. 217
Figure 3.10 RNA-Seq identifies a putative bis-MGD-dependent dopamine dehydroxylase (Dadh) in *E. lenta* A2 .............................................................. 221
Figure 3.11 Model for tungstate inhibition of dopamine dehydroxylation in *E. lenta* A2 .... 222
Figure 3.12 Tungstate inhibits dopamine dehydroxylation without inhibiting growth of *E. lenta* A2.............................................................................................................................................. 223

Figure 3.13 Tungstate does not inhibit dopamine dehydroxylation in cell lysates............. 224

Figure 3.14 Predicted model for *E. lenta* A2 Dadh. ................................................................. 226

Figure 3.15 Dopamine dehydroxylase activity is not membrane-associated in *E. lenta* lysates. 227

Figure 3.16 Ammonium sulfate precipitation of dopamine dehydroxylase activity. ............ 228

Figure 3.17 Native purification of the dopamine dehydroxylase from *E. lenta* A2. ............... 229

Figure 3.18 Dadh-containing fractions purified from *E. lenta* A2 regiospecifically dehydroxylate dopamine to produce *m*-tyramine. ........................................................................................................... 231

Figure 3.19 Global alignment of the dopamine dehydroxylase locus across an Actinobacterial library.................................................................................................................................................. 235

Figure 3.20 A single amino acid variant predicts dopamine metabolism by *E. lenta* and related strains and does not correlate with phylogeny................................................................. 237

Figure 3.21 Alignment of Dadh protein sequences from the Actinobacterial strain library...... 239

Figure 3.22 The Dadh SNP predicts dopamine dehydroxylation by complex gut microbiota samples from neurologically healthy patients ex vivo................................................................. 241

Figure 3.23 Metagenomic analysis of *E. lenta/dadh* and *E. faecalis/tyrDC* abundance and prevalence across human patients.............................................................................................................. 242

Figure 3.24 SNP analysis of *dadh* in human gut metagenomes containing high-coverage *E. lenta* genomes. .............................................................................................................................................. 243

Figure 3.25 The Dadh SNP predicts dopamine dehydroxylation by complex gut microbiota samples Parkinson’s disease patients ex vivo.................................................................................. 244

Figure 3.26 *E. faecalis* and *E. lenta* produce *m*-tyramine from L-dopa in co-culture.......... 246
Figure 3.27 Metabolism of L-dopa by co-cultures of *E. faecalis* and *E. lenta* strains in the presence of AFMT. ................................................................. 247
Figure 3.28 Gain-of-function studies in stool samples incubated with *d*<sub>3</sub>-phenyl-L-dopa. ............ 248
Figure 3.29 Anaerobic growth of *E. lenta* A2 in the presence of carbidopa or AFMT. ............. 249
Figure 3.30 Dehydroxylolation of dopamine by *E. lenta* A2 in the presence of carbidopa or AFMT. ........................................................................................................ 250
Figure 3.31 Predicted metabolites in gnotobiotic mice colonized with *E. faecalis* and *E. lenta* and dosed with *d*<sub>3</sub>-phenyl-L-dopa........................................................................................................ 252
Figure 3.32 Levels of *E. faecalis* and *E. lenta* in gnotobiotic mice............................................. 252
Figure 3.33 Urinary recovery and total catechol excretion in gnotobiotic experiment. ............ 254
Figure 3.34 Metabolites detected in gnotobiotic mouse urine subjected to enzymatic hydrolysis prior to extraction........................................................................................................ 255
Figure 3.35 Metabolites detected in gnotobiotic mouse urine subjected to acid hydrolysis prior to extraction........................................................................................................ 256
Figure 3.36 Metabolites detected in gnotobiotic mouse cecal contents subjected to enzymatic hydrolysis prior to extraction................................................................. 258
Figure 3.37 Metabolites detected in gnotobiotic mouse stool samples subjected to enzymatic hydrolysis prior to extraction................................................................. 259
Figure 3.38 Metabolites detected in gnotobiotic mouse stool samples not subjected to enzymatic hydrolysis prior to extraction................................................................. 260
Figure 3.39 Metabolites detected in gnotobiotic mouse ileal contents......................................... 260
Figure 4.1 Classes of catechol compounds encountered by the human gut microbiota. ........ 308
Figure 4.2 Examples of catechol dehydroxylation reactions performed by the human gut microbiota. .......................................................................................................................... 310

Figure 4.3 SDS-PAGE of natively purified dopamine dehydroxylase from *E. lenta* A2......... 316

Figure 4.4 Dadh specifically metabolizes catecholamines and does not dehydroxylate other physiologically relevant catechol substrates................................................................. 317

Figure 4.5 Dadh tolerates only minor changes to the catecholamine scaffold. ...................... 319

Figure 4.6 The catecholamine scaffold is required for *dadh* induction and dehydroxylation by *E. lenta* A2 cells. ..................................................................................................................... 321

Figure 4.7 Dopamine does not promote growth of *E. lenta* A2 in BHI medium............... 324

Figure 4.8 Dopamine specifically promotes growth of dopamine-dehydroxylating *E. lenta* strains in basal medium. ........................................................................................................ 324

Figure 4.9 Dopamine promotes growth of *E. lenta* A2 in basal medium in a dose-dependent manner............................................................................................................................. 325

Figure 4.10 Tungstate inhibits dopamine-dependent growth and dopamine metabolism by *E. lenta* A2 grown in basal medium................................................................. 326

Figure 4.11 Tungstate and molybdate alone do not impact *E. lenta* A2 growth in basal medium. .............................................................................................................................. 327

Figure 4.12 The dopamine metabolizer *E. lenta* A2 outcompetes the non-metabolizer *E. lenta* DSM2243 in the presence of dopamine in basal medium. ........................................ 328

Figure 4.13 The dopamine metabolizer *E. lenta* Valencia outcompetes the non-metabolizer *E. lenta* W1BHI6 in the presence of dopamine in basal medium...................................................... 329

Figure 4.14 Dopamine promotes growth of metabolizing *E. lenta* strains in basal medium in the presence of a defined microbial community.............................................................. 332
Figure 4.15 Dopamine increases *dadh* and *E. lenta* abundance in complex human gut microbiotas cultured in basal medium ex vivo. .......................................................... 333
Figure 4.16 Dopamine increases the proportion of metabolizing *dadh* variants in complex human gut microbiotas cultured in basal medium ex vivo. .......................................................... 334
Figure 4.17 Colorimetric screen for catechol dehydroxylation by human gut Actinobacteria... 339
Figure 4.18 Confirmation of catechol dehydroxylation by *E. lenta* strains................................. 341
Figure 4.19 Confirmation of catechol dehydroxylation by *G. pamelaeae*. .............................. 343
Figure 4.20 Gut Actinobacteria display strain-level variability in the dehydroxylation of DOPAC, hydrocaffeic acid, and (+)-catechin. .......................................................... 345
Figure 4.21 Hydrocaffeic acid and (+)-catechin promotes growth of *E. lenta* A2 in basal medium. ................................................................................................................. 347
Figure 4.22 DOPAC may promote growth of *G. pamelaeae* 3C in the presence of formate in basal medium. ................................................................................................................. 348
Figure 4.23 Catechol dehydroxylation by whole cell suspensions of *E. lenta* A2 and *G. pamelaeae* 3C is inducible and oxygen sensitive................................................................. 350
Figure 4.24 Specificity and inducibility of gut Actinobacterial dehydroxylation as assessed by activity of cell lysates................................................................. 352
Figure 4.25 Tungstate does not inhibit *E. lenta* A2 growth in BHI medium......................... 361
Figure 4.26 Impact of tungstate on catechol metabolism by *E. lenta* A2 and *G. pamelaeae* 3C. ................................................................................................................. 362
Figure 4.27 Candidate catechol dehydroxylases and their associated metabolic activities are variably distributed across the gut Actinobacterial library used in our study................. 363
Figure 4.28 SDS-PAGE of partially purified hydrocaffeic acid dehydroxylase from E. lenta A2. .......................................................... 365

Figure 4.29 In vitro activity of Hcdh-containing fractions purified from E. lenta A2. ............ 366

Figure 4.30 Peptide coverage of the hydrocaffeic acid dehydroxylase from E. lenta A2 from proteomics................................. 367

Figure 4.31 E. lenta A2 Hcdh displays a narrow substrate scope and requires the hydrocaffeic acid core scaffold for metabolism. .......................................................... 368

Figure 4.32 Time course of DOPAC dehydroxylation by G. pamelaeae 3C. ................. 371

Figure 4.33 Lysate assays suggest the presence of a G. pamelaeae 3C hydrocaffeic acid dehydroxylase distinct from E. lenta Hcdh.......................................................... 381

Figure 4.34 Gordonibacter and Eggerthella dehydroxylases display differences in genomic contexts and predicted subunit compositions. .......................................................... 387

Figure 4.35 Catechol dehydroxylases display variable prevalence and abundance in human gut metagenomes.......................................................... 390

Figure 4.36 Human gut microbiota communities display variable metabolism of catechols ex vivo. .................................................................................................................. 392

Figure 4.37 dodh abundance correlates with DOPAC metabolism by complex human gut communities ex vivo. .......................................................... 393

Figure 4.38 Catechol dehydroxylases are distinct from other bis-MGD enzymes............. 400

Figure 4.39 Sequence similarity network of the bis-MGD enzyme family reveals that the Gordonibacter and Eggerthella dehydroxylases belong to distinct, uncharacterized clusters... 401
Figure 4.40 Sequence similarity network of the bis-MGD enzyme family reveals that *Gordonibacter* and *Eggerthella* bis-MGD enzymes are widely distributed across the enzyme family. ................................................................. 402

Figure 4.41 GNN analysis of *Gordonibacter* and *Eggerthella* type dehydroxylases........ 405

Figure 4.42 GNN-predicted average operonic architecture of *Gordonibacter* and *Eggerthella* type dehydroxylases................................................................. 407

Figure 4.43 Catechol dehydroxylases are widely distributed across sequenced microbes....... 409

Figure 4.44 Number of catechol dehydroxylase homologs among gut Actinobacterial isolates used in our study. ........................................................................................................ 411

Figure 4.45 Phylogenetic analysis of representative catechol dehydroxylase homologs highlighting sequences included in further analyses......................................................... 414

Figure 4.46 Phylogenetic analysis of putative catechol dehydroxylase sequences from diverse bacterial phyla reveals a relationship with the biochemically characterized molybdenum-dependent enzyme acetylene hydratase........................................................................ 415

Figure 4.47 A putative catechol dehydroxylase is critical for the biosynthesis of the anti-tumor compound CC-1065................................................................. 417

Figure 4.48 A putative catechol dehydroxylase is important for *T. aromatica* AR-1 growth on 3,5-dihydroxybenzoic acid................................................................. 418

Figure 4.49 Screen for catechol dehydroxylation by gut microbiota samples from diverse mammals................................................................. 420

Figure 4.50 Gut microbiotas of mammals representing distinct diets and phylogenetic origins can dehydroxylate catechols................................................................. 422

Figure 4.51 Host biosynthesis and gut Actinobacterial metabolism of catecholamines........ 424
Figure 4.52 Preparation of 2-phenylacetic acid derivatives.......................................................... 472
Figure 4.53 Reduction of benzyl carboxylic acids........................................................................ 472
Figure 4.54 Chemical structure of dimethoxyphenethyl alcohol (S1).......................................... 473
Figure 4.55 Chemical structure of dimethoxyphenethyl alcohol (S2).......................................... 473
Figure 4.56 Chemical structure of dimethoxyphenethyl alcohol (S3).......................................... 474
Figure 4.57 Chemical structure of trimethoxyphenethyl alcohol (S4)........................................ 474
Figure 4.58 Chemical structure of trimethoxyphenethyl alcohol (S5)........................................ 475
Figure 4.59 Scheme for Mitsunobu reaction on phenethyl alcohols............................................. 475
Figure 4.60 Chemical structure of dimethoxyphenethyl phthalimide (S6)................................. 475
Figure 4.61 Chemical structure of Dimethoxyphenethyl phthalimide (S7)................................. 476
Figure 4.62 Chemical structure of imethoxyphenethyl phthalimide (S8).................................... 476
Figure 4.63 Chemical structure of trimethoxyphenethyl phthalimide (S9).................................. 477
Figure 4.64 Chemical structure of trimethoxyphenethyl phthalimide (S10)............................... 478
Figure 4.65 Hydrazinolysis of phenethyl phthalimides................................................................. 478
Figure 4.66 Chemical structure of dimethoxyphenethyl amine (S11)......................................... 479
Figure 4.67 Chemical structure of dimethoxyphenethyl amine (S12)......................................... 479
Figure 4.68 Chemical structure of dimethoxyphenethyl amine (S13)......................................... 480
Figure 4.69 Chemical structure of trimethoxyphenethyl amine (S14)........................................ 480
Figure 4.70 Chemical structure of trimethoxyphenethyl amine (S15)......................................... 480
Figure 4.71 Practical demethylation of aryl methyl ethers............................................................ 481
Figure 4.72 Chemical structure of aminoethylbenzene diol hydrobromide (S16)...................... 481
Figure 4.73 Chemical structure of aminoethylbenzene diol hydrobromide (S17)...................... 482
Figure 4.74 Chemical structure of 5-(2-Aminoethyl)benzene-1,4-diol hydrobromide (S18). .. 482
Figure 4.75 Chemical structure of aminoethylbenzene triol hydrobromide (S19). ....................... 483
Figure 4.76 Chemical structure of aminoethylbenzene triol hydrobromide (S20). ....................... 483
Figure 4.77 Preparation of aminoethylbenzene triol hydrobromide from aryl nitriles .......... 484
Figure 4.78 Chemical structure of trimethoxybenzonitrile (S21). ............................................. 485
Figure 4.79 Chemical structure of trimethoxybenzaldehyde (S22). ........................................ 485
Figure 4.80 Chemical structure of nitrovinyl benzene (S23). .................................................... 486
Figure 4.81 Chemical structure of nitrovinyl benzene (S23). .................................................... 487
Figure 4.82 Chemical structure of aminoethylbenzene triol hydrobromide (S25). ............... 487
Figure 4.83 Preparation of aminoethylbenzene triol hydrobromide derivatives from trimethoxybenzene .............................................................. 488
Figure 4.84 Chemical structure of trimethoxy nitrovinyl benzene (S26). ............................... 489
Figure 4.85 Chemical structure of trimethoxyphenethyl amine (S27). ................................. 489
Figure 4.86 Chemical structure of trimethoxyphenethyl amine (S27). .................................. 490
Figure 4.87 2-Amino-4-(2-aminoethyl)phenol via Mitsunobu strategy. ................................. 491
Figure 4.88 Chemical structure of benzyloxy nitrophenethyl alcohol (S29). .......................... 491
Figure 4.89 Chemical structure of benzyloxy nitrophenethyl phthalimide (S30). ..................... 492
Figure 4.90 Chemical structure of amino alcohol (S31). .......................................................... 493
Figure 4.91 Chemical structure of 2-Amino-4-(2-aminoethyl)phenol (S32). ........................ 494
Figure 4.92 Reduction of a dihydroxyphenyl acetic acid. ....................................................... 495
Figure 4.93 ¹H NMR of dimethoxyphenethyl alcohol (S1). ..................................................... 497
Figure 4.94 ¹³C NMR of dimethoxyphenethyl alcohol (S1). ..................................................... 498
Figure 4.95 ¹H NMR of trimethoxyphenethyl alcohol (S4). ..................................................... 499
Figure 4.96 ¹³C NMR of trimethoxyphenethyl alcohol (S4). ..................................................... 500
Figure 4.97 $^1$H NMR of trimethoxyphenethyl alcohol (S5). ................................................................. 501
Figure 4.98 $^{13}$C NMR of trimethoxyphenethyl alcohol (S5). ................................................................. 502
Figure 4.99 $^1$H NMR of trimethoxyphenethyl phthalimide (S10). ............................................................. 503
Figure 4.100 $^{13}$C NMR of trimethoxyphenethyl phthalimide (S10). ............................................................. 504
Figure 4.101 $^1$H NMR of dimethoxyphenethyl amine (S11). ................................................................. 505
Figure 4.102 $^{13}$C NMR of dimethoxyphenethyl amine (S11). ................................................................. 506
Figure 4.103 $^1$H NMR of dimethoxyphenethyl amine (S12). ................................................................. 507
Figure 4.104 $^{13}$C NMR of dimethoxyphenethyl amine (S12). ................................................................. 508
Figure 4.105 $^1$H NMR of trimethoxyphenethyl amine (S14). ................................................................. 509
Figure 4.106 $^{13}$C NMR of trimethoxyphenethyl amine (S14). ................................................................. 510
Figure 4.107 $^1$H NMR of trimethoxyphenethyl amine (S15). ................................................................. 511
Figure 4.108 $^{13}$C NMR of trimethoxyphenethyl amine (S15). ................................................................. 512
Figure 4.109 $^1$H NMR of aminoethylbenzene diol hydrobromide (S16). ..................................................... 513
Figure 4.110 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S16). ..................................................... 514
Figure 4.111 $^1$H NMR of aminoethylbenzene diol hydrobromide (S17). ..................................................... 515
Figure 4.112 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S17). ..................................................... 516
Figure 4.113 $^1$H NMR of aminoethylbenzene diol hydrobromide (S18). ..................................................... 517
Figure 4.114 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S18). ..................................................... 518
Figure 4.115 $^1$H NMR of aminoethylbenzene triol hydrobromide (S19). ..................................................... 519
Figure 4.116 $^{13}$C NMR of aminoethylbenzene triol hydrobromide (S19). ..................................................... 520
Figure 4.117 $^1$H NMR of nitrovinyl benzene (S23). ................................................................. 521
Figure 4.118 $^{13}$C NMR of nitrovinyl benzene (S23). ................................................................. 522
Figure 4.119 $^1$H NMR of trimethoxyphenethyl amine (S24). ................................................................. 523
Figure 4.120 $^{13}$C NMR of trimethoxyphenethyl amine (S24). .......................................................... 524
Figure 4.121 $^1$H NMR of aminoethylbenzene triol hydrobromide (S25). .............................................. 525
Figure 4.122 $^{13}$C NMR of aminoethylbenzene triol hydrobromide (S25). ............................................. 526
Figure 4.123 $^1$H NMR of trimethoxy nitrovinyl benzene (S26). ............................................................. 527
Figure 4.124 $^{13}$C NMR of trimethoxy nitrovinyl benzene (S26). ............................................................. 528
Figure 4.125 $^1$H NMR of trimethoxyphenethyl amine (S27). ............................................................... 529
Figure 4.126 $^{13}$C NMR of trimethoxyphenethyl amine (S27). ............................................................... 530
Figure 4.127 $^1$H NMR of benzyloxy nitrophennethyl alcohol (S29).......................................................... 531
Figure 4.128 $^{13}$C NMR of benzyloxy nitrophennethyl alcohol (S29).......................................................... 532
Figure 4.129 $^1$H NMR of benzyloxy nitrophennethyl phthalimide (S30). ................................................. 533
Figure 4.130 $^{13}$C NMR of benzyloxy nitrophennethyl phthalimide (S30). ................................................. 534
Figure 4.131 $^1$H NMR of amino alcohol (S31). ............................................................................................ 535
Figure 4.132 $^{13}$C NMR of amino alcohol (S31). ............................................................................................ 536
Figure 4.133 $^1$H NMR of 2-amino-4-(2-aminoethyl)phenol (S32). ......................................................... 537
Figure 4.134 $^{13}$C NMR of 2-amino-4-(2-aminoethyl)phenol (S32). ......................................................... 538
Figure 5.1 The molybdenum-dependent nitrogenase harbors a heteronuclear cofactor. ........... 555
Figure 5.2 Mononuclear molybdenum-dependent enzymes. ................................................................. 557
Figure 5.3 The bis-MGD cofactor and common reactions catalyzed by family members....... 559
Figure 5.4 Proposed mechanism for DMSO reduction by R. sphaeroides DMSO reductase. ... 561
Figure 5.5 Proposed mechanism of hydroxy transfer by pyrogallol hydroxytransferase........... 564
Figure 5.6 Proposed mechanism of acetylene hydration by acetylene hydratase...................... 567
Figure 5.7 Proposed mechanism for catechol dehydroxylation by catechol dehydroxylases.... 570
Figure 5.8 Biosynthesis of the bis-MGD cofactor. ................................................................................. 572
Figure 5.9 Subunit composition and cellular location of microbial bis-MGD enzymes. ........... 576
Figure 5.10 Lysate assays reveal that hydrocaffeic acid dehydroxylation incorporates high levels of a single deuterium into the product m-hppa ................................................................. 587
Figure 5.11 High-resolution LC-MS/MS indicates that a single deuterium is specifically incorporated into the aromatic ring of m-hppa. ................................................................. 588
Figure 5.12 Analysis of dadh and its surrounding genomic context. .................................. 590
Figure 5.13 Aerobic and anaerobic expression of DadhA in E. coli BL21 (DE3) in LB medium supplemented with nitrate, molybdate, and ferric ammonium citrate. ................................. 593
Figure 5.14 Aerobic and anaerobic expression of DadhA in E. coli BL21 (DE3) harboring the pPH149 plasmid in LB medium supplemented with nitrate, cysteine, molybdate, and ferric ammonium citrate. .................................................................................................................. 595
Figure 5.15 Anaerobic expression of selected DadhA constructs in E. coli BL21 (DE3) adapting previous protocols for acetylene hydratase and arsenate reductase expression .................. 597
Figure 5.16 Attempted purification of C-terminally His6-tagged full-length DadhA from E. coli BL21 (DE3). ................................................................................................................................. 599
Figure 5.17 Multiple gene expression of Dadh in E. coli BL21 (DE3) ΔiscR and E. coli MG1655. ..................................................................................................................................................... 603
Figure 5.18 Expression of DadhABC on pTrcHis2A in E. coli MG1655. ......................... 604
Figure 5.19 Expression of DadhABC on pTrcHis2A in with and without auxiliary plasmids in E. coli MG1655 and E. coli K-12 ΔiscR. .......................................................................................... 605
Figure 5.20 Expression of DadhABCE on pETDuet1 in E. coli MG1655(DE3) ............... 606
Figure 5.21 Predicted subunit composition of Dadh homologs selected for heterologous expression. ............................................................................................................................... 610
Figure 6.1 Implications of understanding gut microbial metabolism of xenobiotics. ............... 663

Figure 6.2 Potential substrates for uncharacterized dehydroxylases. .................................. 674
List of Tables

Table 2.1 Michaelis–Menten parameters determined for TyrDC. .................................................. 135
Table 2.2 Inhibition parameters determined for carbidopa, benzerazide, and benzerazide free hydrazine ......................................................................................................................................... 155
Table 2.3 Inhibition parameters determined for AFMT. .................................................................. 162
Table 2.4 Genomic locations of transposon mutants in E. faecalis MMH594. .............................. 175
Table 3.1 Genes upregulated in E. lenta A2 in response to dopamine. ........................................... 220
Table 3.2 Proteomics identification of Dadh. .................................................................................. 231
Table 3.3 Predicted dopamine dehydroxylases among sequenced isolates. ...................................... 234
Table 3.4 Dopamine dehydroxylases in gut Actinobacteria. ............................................................. 236
Table 3.5 Genes upregulated in E. lenta DSM2243 when grown with 500 µM dopamine.... 238
Table 3.6 Genes upregulated in E. lenta 28b when grown in the presence of dopamine............. 238
Table 4.1 Physiologically relevant catechol substrates used in assays with natively purified dopamine dehydroxylase from E. lenta A2. ......................................................................................... 315
Table 4.2 Commercially available and synthesized dopamine analogs used in assays with natively purified dopamine dehydroxylase from E. lenta A2............................................................ 318
Table 4.3 Genes differentially expressed upon exposure of E. lenta A2 to norepinephrine. ...... 322
Table 4.4 Bacterial strains used in this study............................................................................... 331
Table 4.5 Catechol substrates used in screen of gut Actinobacteria for catechol metabolism. .... 337
Table 4.6 Genes differentially expressed upon exposure of Eggerthella lenta A2 to hydrocaffeic acid...................................................................................................................................... 354
Table 4.7 Genes differentially expressed genes upon exposure of Eggerthella lenta A2 to (+)-catechin. ................................................................................................................................. 356
Table 4.8 Percent amino acid identity between the putative catechol dehydroxylases from *G. pamelaeae* 3C and *E. lenta* A2. .................................................................................................................. 359

Table 4.9 Proteomics identification of proteins in the band of interest during the activity-based native purification of the hydrocaffeic acid dehydroxylase from *E. lenta* A2................................. 367

Table 4.10 Genes differentially expressed upon exposure of *G. pamelaeae* 3C to DOPAC when the catechol is added during exponential phase.................................................................................................................. 372

Table 4.11 Genes differentially expressed upon exposure of *G. pamelaeae* 3C to DOPAC when the catechol is added at the time of inoculation.................................................................................................................. 377

Table 4.12 Genes differentially expressed upon exposure of *E. lenta* A2 to caffeic acid. ........ 383

Table 4.13 Genes differentially expressed upon exposure of *E. lenta* 28b to norepinephrine. .. 384

Table 4.14 Accession numbers (Uniprot and Genbank) of putative *Eggerthella* and *Gordonibacter* dehydroxylases identified in this study. .................................................................................................................. 386

Table 4.15 Accession numbers of bis-MGD enzymes used to generate the phylogenetic tree of the bis-MGD enzyme family (Figure 4.40). .................................................................................................................. 396

Table 4.16 Accession numbers (Uniprot) and origin of representative dehydroxylases identified from phylogenetic tree in Figure 4.43.................................................................................................................. 413

Table 5.1 Dadh constructs expressed on pET vectors. ................................................................................. 592

Table 5.2 Strategy for Dadh multiple gene expression in *E. coli*. ........................................................... 603

Table 5.3 Plasmids used for expression of DadhABC on pTrcHis2A in with and without auxiliary plasmids in *E. coli* MG1655 and *E. coli* K-12 ΔiscR.................................................................................................. 605

Table 5.4 Plasmids used for multiple gene expression in *E. coli* MG1655 (DE3). ...................... 606

Table 5.5 Plasmids yet to be co-expressed in *E. coli* MG1655. ................................................................. 607

Table 5.6 Plasmids expressed in *R. erythropolis* L-88 and *E. faecalis* OG1RF. .............................. 608
Table 5.7 Expression of dehydroxylase homologs in *Shewanella sp.* ANA-3 and *E. coli.* ...... 612

Table 5.8 Expression of Dodh in *E. coli.* .................................................................................................. 613
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>° C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>2,3-DHBA</td>
<td>2,3-dihydroxybenzoic acid</td>
</tr>
<tr>
<td>3,4-DHMA</td>
<td>3,4-dihydroxymandelic acid</td>
</tr>
<tr>
<td>3,5-DHB</td>
<td>3,5-dihydroxybenzoic acid</td>
</tr>
<tr>
<td>4-HBCR</td>
<td>4-hydroxybenzoyl-CoA reductase</td>
</tr>
<tr>
<td>5-ASA</td>
<td>5-Aminosalicylic acid</td>
</tr>
<tr>
<td>5-FU</td>
<td>5-fluorouracil</td>
</tr>
<tr>
<td>AADC</td>
<td>aromatic amino acid decarboxylase</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>AFMT</td>
<td>(S)-α-fluoromethyltyrosine</td>
</tr>
<tr>
<td>Arg, R</td>
<td>Arginine</td>
</tr>
<tr>
<td>Aro</td>
<td>arsenite oxidase</td>
</tr>
<tr>
<td>Arr</td>
<td>arsenate reductase</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BHI</td>
<td>brain heart infusion media</td>
</tr>
<tr>
<td>BRV</td>
<td>Brivudine</td>
</tr>
<tr>
<td>BVU</td>
<td>(E)-5-(2-bromovinyl)-uracil</td>
</tr>
<tr>
<td>Cadh</td>
<td>catechin dehydroxylase</td>
</tr>
<tr>
<td>CD</td>
<td>celiac disease</td>
</tr>
<tr>
<td>Cdh</td>
<td>catechol dehydroxylase</td>
</tr>
<tr>
<td>CFU</td>
<td>colony forming unit</td>
</tr>
</tbody>
</table>
CGFP  chemically guided functional profiling
Cldh  catechol lignan dehydroxylase
COMT  Catechol-O-Methyltransferase
cPMP  cyclic pyranopterin monophosphate
CPT-11  Irinotecan
CutC  choline trimethylamine-lyase
Cys, C  cysteine
DA  dopamine
Dadh  dopamine dehydroxylase
DBH  dopamine-beta hydroxylase
DMF  dimethyl formamide
Dms  DMSO reductase
dmSECO  didemethylsecoisolariciresinol
DMSO  dimethyl sulfoxide
Dodh  dopac dehydroxylase
DOPAC  3,4-dihydroxyphenylacetic acid
DPD  dihydropyrimidine dehydrogenase
Ebd  ethylbenzene dehydrogenase
EIC  Extracted Ion Chromatogram
El  *E. lenta*
EPR  electron paramagnetic resonance
Es  *E. sinensis*
FAD  flavin adenine dinucleotide
FC           fold-change
Fdh          formate dehydrogenase
FMO          Flavin monooxygenase
FPLC         fast protein liquid chromatography
g            gram
GABA         γ-aminobutyric acid
gDNA         genomic DNA
GI tract      gastrointestinal tract
Gp            Gordonibacter pamelaeae
GRE          glycyl radical enzyme
Gs            Gordonibacter sp
GTP          guanine triphosphate
GUS          β-glucuronidases
h            hour
h            hours
Hcdh         hydrocaffeic acid dehydroxylase
His6         Hexahistidine affinity tag
HMP          Human Microbiome Project
HVA          homovanillic acid
IBD          inflammatory bowel disease
IPTG         isopropyl β-D-1-thiogalactopyranoside
IQ            2-amino-3-methylimidazo[4,5-f]quinoline
ITC          isothiocyanate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>L-dopa</td>
<td>Levodopa</td>
</tr>
<tr>
<td>LB medium</td>
<td>Luria-Bertani medium</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatography-Mass Spectrometry</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>liquid chromatography tandem mass spectrometry</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>m-hpaa</td>
<td>m-hydroxyphenylacetic acid</td>
</tr>
<tr>
<td>m-hppa</td>
<td>m-hydroxyphenylpropionic acid</td>
</tr>
<tr>
<td>MAO</td>
<td>monoamine oxidase</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>Mo</td>
<td>molybdenum</td>
</tr>
<tr>
<td>Moco, bis-MGD</td>
<td>bis-molybdopterin guanine dinucleotide cofactor</td>
</tr>
<tr>
<td>MPT</td>
<td>molybdopterin</td>
</tr>
<tr>
<td>MRM</td>
<td>Multiple reaction monitoring</td>
</tr>
<tr>
<td>MRP</td>
<td>Maillard reaction product</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NaDT</td>
<td>sodium dithionite</td>
</tr>
<tr>
<td>Nar</td>
<td>nitrate reductase</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NZ</td>
<td>nitrazepam</td>
</tr>
<tr>
<td>O-dma</td>
<td>O-desmethylandolensin</td>
</tr>
</tbody>
</table>

xlv
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>$p$-hpaa</td>
<td>$p$-hydroxyphenylacetic acid</td>
</tr>
<tr>
<td>$p$-hppa</td>
<td>$p$-hydroxyphenylpropionic acid</td>
</tr>
<tr>
<td>Pcr</td>
<td>perchlorate reductase</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>Phet</td>
<td>pyrogallol hydroxytransferase</td>
</tr>
<tr>
<td>Ach</td>
<td>acetylene hydratase</td>
</tr>
<tr>
<td>PL</td>
<td>polysaccharide lyase</td>
</tr>
<tr>
<td>PLD</td>
<td>phospholipase D</td>
</tr>
<tr>
<td>PLP</td>
<td>Pyridoxal-L-phosphate</td>
</tr>
<tr>
<td>PNMT</td>
<td>phenylethanolamine N-methyltransferase</td>
</tr>
<tr>
<td>Psr</td>
<td>polysulfide reductase</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>RPM</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEM</td>
<td>standard error of the mean</td>
</tr>
<tr>
<td>Ser, S</td>
<td>Serine</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSN</td>
<td>sequence similarity network</td>
</tr>
<tr>
<td>SSRI</td>
<td>selective serotonin reuptake inhibitor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Sus</td>
<td>starch uptake system</td>
</tr>
<tr>
<td>TyrDC</td>
<td>tyrosine dehydroxylase</td>
</tr>
<tr>
<td>Tardh</td>
<td>T. aromatica dehydroxylase</td>
</tr>
<tr>
<td>TAT</td>
<td>Twin-Arginine-Translocation</td>
</tr>
<tr>
<td>Tet</td>
<td>tetracycline</td>
</tr>
<tr>
<td>TetR</td>
<td>tetracycline resistance</td>
</tr>
<tr>
<td>TetS</td>
<td>tetracycline sensitive</td>
</tr>
<tr>
<td>TMA</td>
<td>trimethylamine</td>
</tr>
<tr>
<td>TMAO</td>
<td>trimethylamine N-oxide</td>
</tr>
<tr>
<td>Tmr</td>
<td>TMAO reductase</td>
</tr>
<tr>
<td>Tris</td>
<td>2-amino-2-(hydroxymethyl)propane-1,3-diol</td>
</tr>
<tr>
<td>TYH</td>
<td>tyrosine hydroxylase</td>
</tr>
<tr>
<td>Tyr, Y</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UV-vis</td>
<td>ultraviolet visible</td>
</tr>
<tr>
<td>VESPR</td>
<td>valence shell electron pair repulsion</td>
</tr>
<tr>
<td>W</td>
<td>tungsten</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
</tbody>
</table>
Acknowledgements

First, I would like to thank my advisor, Emily Balskus. Thank you for inviting me into your lab and into the wonderful world of microbial chemistry. I first came to your research group in 2014 as a summer student in the SROH program hosted by Harvard MCB. At the end of my internship, you joined the final SROH dinner and we talked about the microbiome for hours. I learned that night, among other things, that methanogens were considered a critical energy sink in the human gut. Facts like these intrigued me – little did I know that they would represent the beginning of my journey as a graduate student. Your scientific support and guidance have empowered me as a scientist and allowed me to find my voice as a scholar. Thank you for believing in my research ideas and providing the resources for me to explore them. Thank you for supporting my professional development as I forged (and still try to figure out) my path at the intersection of food and science. Your work ethic and critical approach to research has set examples for breaking boundaries in science and academia.

I also want to acknowledge my mentors at Harvard. Rachelle Gaudet, thank you for your emotional and scientific support. I always emerge from our committee meetings and coffee chats with new perspectives. Your advice was especially critical in helping me navigate my identity as a scientist and decide on my next professional steps. Rich Losick, thank you for serving as my mentor. From informal chats in your office to committee meetings, your feedback on my research and my professional aspirations has been invaluable. Rachel Carmody, thank you for being a friend, mentor, and a scientific collaborator. You have shown me how to create room for food and cooking in research while maintaining a commitment to mentoring relationships and empowering the next generation. I will never forget going to Whole Foods together loading up our two carts for our microbiome cooking study. I am forever grateful for the opportunity you provided to formally
combine my passions for science and cooking in an academic research. Thank you to Rachelle, Rich, and Rachel, for serving on my DAC all these years. Thank you also to Christina Woo and Karine Gibbs for serving on my dissertation committee. In addition, I want to thank Sheila Thomas for supporting my development in the SROH program and supporting my transition to Harvard. Your mentorship was a key ingredient in making me comfortable with my background and confident in my abilities before and during graduate school. I especially appreciate the opportunities you provided to integrate cooking and science education into my graduate student experience. I loved my experiences with the HILS cooking challenge and with the NAHSSP program. Finally, thank you to Kathryn Hollar and Pia Sorensen for inviting me into the science and cooking community that I had dreamed of since early college. Working with you in the Harvard SPU27 class and in the community with teachers and students was a dream come true. You empowered me as an educator and inspired my visions for the future. STEAMed in particular was an incredible experience I will never forget.

I also want to thank the incredible mentors I met outside of Harvard, who laid the foundation of my academic aspirations and achievement, and have provided essential personal, academic, and professional support before and during graduate school. Peter Balaam, thank you for teaching me how to write, for looking beyond my initial failures in your class and pushing me to reach my potential, making me believe that I not only belonged in college but that I could succeed in anything I put my mind to. Thank you for lending a listening ear, helping me navigate personal and financial stress during early college, and for being a supportive and inspiring friend in the complexities of adulthood. Liz Ciner, thank you for showing me the way and guiding me to the destination. You told me about research, about fellowships, and empowered me to seek out and succeed in opportunities I did not even know existed. From working with you I learned not
only how to communicate effectively, but also what it means to be a true mentor. You showed me that sometimes mentorship is the missing link between scientific interest and academic achievement. As I more and more frequently assume the role of mentor in my daily life, I always try to channel my inner Liz, striving for your perfect balance between high challenge and high support. So much of what I have accomplished and learned I owe to you. Deborah Gross, thank you for encouraging me to pursue chemistry, bringing me into the vibrant and supportive department at Carleton, and for providing a space in academia for my interest in food. You convinced me there is room for my passions and for my identity in science and I am forever grateful for your continued mentorship and friendship. Chris Calderone, thank you for inviting me into the wonderful world of biological chemistry. Our chats in your office about enzymes and mechanisms made me feel fuzzy and warm inside and inspired me to pursue a research career in chemical biology. Microbes are amazing! Joe Szursewski and Purna Kashyap at the Mayo Clinic, thank you for inviting me into your labs and to the world of the human gut microbiome. You catalyzed my early interest in research and made me believe I had a place in science. Joe, I distinctly remember one of the first days at Mayo Clinic back in 2013. You brought me into your office and showed me your old X-ray films of barium labels progressing through the dog gastrointestinal tract. From that moment onwards, you always made an active effort to show me the beauty of science and to support me personally and professionally.

Thank you also to my collaborators who have provided invaluable insight and contributions to my research project in graduate school. Thank you to Peter Turnbaugh, Jordan Bisanz, Elizabeth Bess, for inspiring scientific discussions about L-dopa and catechol metabolism and helping with critical experiments. Thank you, Elizabeth Selleck and Michael Gilmore, for sharing your expertise on *Enterococcus* genetics, evolution, and microbiology.
Thank you to the members of the Balskus lab, past and present. You have been great colleagues. Our interactions have catalyzed many new ideas and progress. Thank you for sharing your time and expertise. In particular, thank you to Matt Wilson, my main man and mentor during my early days in the lab and someone I keep learning from still today. Thank you for always sharing your wisdom and experience and providing feedback on my millions of questions, both personal and professional. I am glad I also finally taught you how to cook. Thank you to Nathaniel Braffman, a great colleague, fellow vacation activist, and friend with whom I could navigate the graduate school experience. I am excited for the day when you are writing your acknowledgements. Michael Luescher, I loved our squash games and chatting about everything from soccer to reaction mechanisms. Thank you for your intellectual and practical investment in my project. Walter Sandoval, I would not have survived without you, knowing I had someone from a cultural and personal background similar to my own made all the difference. Chip Le, your arrival in the lab improved my quality of life immensely. It has been such a privilege to work with someone so intelligent and motivated to develop the project into the future. You have shown me the true beauty of collaboration – from our interactions, I have learned much about science, but most importantly that I never want to work completely alone again. Our exchanges and jokes have fueled progress and inspired novel ideas that are greater than the sum of their parts. To my undergraduate Sina Kiamehr, it has been a pleasure to work with you and observe your personal and professional growth. Your enthusiasm and engagement have been a shining light since January 2019. Mentoring someone from a very similar background to my own has taught me many things about science and about what it means to set an example and pay it forward to the next generation. To Minwoo, it has been a privilege to mentor someone so intelligent and hardworking as you. The project could not be in better hands. Your biggest challenge will be to learn how to take a break!
Thank you to Xueyang, I have really enjoyed thinking about new directions for my project and exploring a completely new avenue of research together. Paula Pelayo, thank you for joining the laboratory after we first chatted at the MCB retreat back in 2018. It warms my heart to have been part of your early scientific journey and to have someone to talk about things very few others can relate to. Keep crushing it. Finally, thank you to Chip and Minwoo for going through the more than 650 pages of this thesis. It literally would not be the same without you.

Para Juanita y Gloria, miembros de la personal de limpieza en CCB. Me dieron la oportunidad acceder a mi herencia Latina. mejorado mi espanol, ayudandome sobrevivir cuando trabaje en Pais Vasco. Ha sido un placer conocerles. Voy a echar de menos nuestras conversaciones en Mallickrodt y en nuestras casas. Me han recordado que hay humanidad en lugares que a veces la carecen. Me recuerdan de mi pasado, mi cultura, y mi familia. Me inspiran hacer una diferencia en mi comunidad. Mil gracias.

To the people I have met through the Gilliam fellowship, thank you for the incredible community. You made me more comfortable in my own skin. You made me realize that my cultural and personal identity are not at odds with my identity as a scientist. The friends and lessons I gained from my time as a Gilliam fellow will last with me forever. I am a proud illustration of how programs like Gilliam are critical for increasing diversity in science.

To the MCO and MCB administration, thank you for always being there and showing that you care. Venkatesh Murthy, thank you for taking me under your wings early on, bringing me to the Indian store in Waltham and allowing me to cook for you at my house and share my passion with the MCB292 class back in fall 2015. I have also immensely appreciated your leadership during the COVID-19 crisis. Catherine Dulac, thank you for engaging with my crazy ideas already back at MCO recruitment, reading through fellowship applications, and showing that you cared
about my ideas and development. Thank you to Fanuel Muindi and all the other administrators who have ensured I always have answers when I need them and for making sure the ride was as smooth as possible. You have reminded me of why I chose to come to the MCO program in the first place.

To my friends in graduate school and beyond, thank you. My roommates, Roberto, Aditi: I have loved to come home to such a diverse household full of life and warmth. Our dinners and conversations were much needed during COVID-19 quarantine and the writing of this thesis. Anne-Sophie, I love our frequent lunches and our cookouts at my house. It’s been nothing short of amazing to follow each other’s journeys – from making foraged pumpkin soup at home as roommates, to sharing our struggles over Skype when you did fieldwork in Mongolia. Your energy is infectious. Alan Pacheco and Anthony Ayala, we should soon be eligible for Platinum status at Mala. You have been a frequent source of support and joy, and I have loved to explore new areas and ideas with you two and Paula Pelayo, Fred Moss, Jose Velilla, and other people! No perreo sola. Perreamos juntos. Charlotte Mann, how lucky was I that the best pastry chef I know and my good friend moved to Boston when I needed it the most? You are dearly missed but I can’t wait for the day you are writing your acknowledgements section on your thesis. I will remember our memories cooking at Carleton and in Cambridge forever. Jackson, Callum, thank you for the frequent laughs and regular conversations. I can’t believe you survived three trips to Mala in one weekend. This strength testifies to the strength of our friendship. Joy, Veronica, Samori, Dragan, and all my friends back in Sweden, thank you for keeping me connected and reminding that there is a world outside my worries in academia and science. Staying connected to friends from my past allowed me to stay grounded in the present. Finally, Chris Bakerlee, thank you for everything. Great conversations about life, about our careers – over dinner or over sports. You inspire me and

liii
it is a privilege to call you a friend. And I am glad you found Frank when he escaped in New Hampshire.

Thank you to Edie and Charlie, who have provided a home away from home in Cambridge. Ever since I met you in 2014 I have loved our conversations and our explorations of food. You are all that I could have asked for and have provided a safe haven during times of crisis and joy. Knowing I always had something to fall back on has minimized the pain of being away from my own family and allowed me to focus on the task at hand, including my dissertation. Soon I will be driving you around!

To my family, thank you. This accomplishment is as much yours as it is mine. I was raised by a community and you have all influenced my character, passions, and aspirations. To my grandparents, thank you for forging your paths in new, unfamiliar countries, overcoming cultural, financial, and personal challenges to lay the foundation for me to succeed. To my mother, Sujata, thank you for introducing me to the kitchen, to the wonders of flavor and texture, to the wonders of the natural world. I found my way to science from our early kitchen experiments. Your love for food and for me and my sister has been my fuel since day one. Thank you for opening my eyes to what it means to be inclusive what it means to be good. Your incredible work as a single parent reminds me that no hill is too large to climb. It also taught me how to hustle and be creative in the lab and beyond. To Yuvia, my sister. If I ever feel lost, I just need to look at you. You are my witness, my lifeline. You remind me of who we are and where we came from. Being your brother and having you as a friend is one of my biggest joys in life and is a main reason I could reach for the stars even when the sky looked cloudy. Oceans apart you are always in my heart. Your creativity inspires me every day. To Radha, my aunt, thank you. I have loved your visits, our sports practice and conversations. Your kindness and generosity have been essential from early childhood.
to adulthood, and I love your engagement with my research and relationships today. Finally, thank you to Johan, my godfather. You are one of the big reasons I am where I am and who I am today. You have shown me that being a real father has nothing to do with blood, that love is the most important foundation of a relationship. Thank you to Janna, your wife, and Pelle, your son and who I call my brother, for showing me that family is just a matter of definition. I know you love sailing, so here is an analogy you will appreciate. You not only kept my sailing boat afloat, but you also provided the navigation tools and wind to keep it on the right course – from day 1 onwards.

Finally, to Tatiana Hill, or Dr. Tatiana Hill as you probably prefer, my partner and best friend, thank you. I came to Harvard as a summer student in 2014 with curiosity and excitement. I left with a newfound passion for science and a connection to you. Ever since, you have revealed to me the beauty of love and of life itself. Your intelligence and passion inspire me, your kindness and optimism comfort me. Whether through a crappy Facetime connection at night or summer bike trips on remote Danish islands, you are an endless source of support and joy. And thank you for always educating me about my ethnic identity and emotional development through the lens of an educational psychologist. The strength of our relationship gave me the strength to believe in myself. Six years apart, our bond is stronger than I could ever dream of. I am privileged to have found you and my new family. I can finally cook for you every single day. We made it.
Chapter 1. Gut microbial metabolism of xenobiotics and endogenous compounds

This chapter was adapted in part from:

1.1 The human microbiota

We live in a microbial world. Microorganisms such as bacteria, archaea, fungi, and viruses colonize virtually every corner of the planet, including our bodies. Humans have co-evolved with trillions of microbes inhabiting the urogenital tract, the skin, the oral cavity, and the gastrointestinal tract. Our interest in these microbial communities (collectively named the human microbiota) dates back to observations made hundreds of years ago by van Leeuwenhoek, but recent advances in DNA sequencing have transformed our understanding of human-associated microbes and their links to health and disease. While changes in microbial community composition are associated with many aspects of human health and disease, it remains unclear in many cases whether the microbiota plays a causative role in host phenotypes or whether these changes merely reflect changes in host physiology. To elucidate the biological consequences of harboring a microbiota, we must move beyond simply knowing what microbes live in and on our body, to understanding what they are doing at a functional level. Dissecting host-microbe and microbe-microbe interactions down to the level of individual strains, genes, and molecules represents a challenge for microbiome studies but is essential for moving from a descriptive to a mechanistic understanding of the human microbiota.

1.2 The human gut microbiota and its involvement in metabolism

The microbial community inhabiting the human gut is one of the most diverse microbial communities in the human body and one of the densest on the planet. The different regions of the human gastrointestinal tract vary in epithelial cell physiology, pH, oxygen levels, and nutrient content, which provides distinct habitats for microorganisms and influences microbial community structure and function (1, 2) (Figure 1.1). For example, only a few organisms can tolerate the harsh conditions of the acidic stomach. The largely aerobic small intestine harbors a more
accommodating pH and a greater density of microbes than the stomach. Moving from the upper small intestine towards the large intestine, the microbial density increases significantly, and in the lower parts of the large intestine, which is largely anaerobic, microbial density can reach $10^{12}$ bacteria/mL. This is the densest microbial habitat in the gastrointestinal tract and a major site of microbial activity. As a whole, the human gut is colonized by hundreds of different microbial species, including some strains not found in other habitats (3, 4). Although obligate anaerobes such as members of the Firmicutes and Bacteroidetes phyla typically dominate, large variability in community composition is observed among individuals and across age and geography (4-8).

Figure 1.1 Map of the human gastrointestinal tract and locations of gut microbes. Microbial estimates are represented as numbers of bacteria per mL of intestinal contents (1).
Possessing 150-fold more genes than the human genome, the human gut microbiota harbors vast metabolic capabilities that expand the range of chemistry taking place in the body (1, 9, 10). Microbial contributions to metabolism were made clear as early as the 1950s, when changes in metabolism in the absence of microbes (for example in germ-free (GF) animals who do not harbor a microbiota) or upon microbial perturbation (i.e., antibiotic treatment or dietary modulation) indicated the gut microbiota’s involvement in specific transformations (11, 12). Further work has illuminated critical biological roles for these reactions; microbial metabolism breaks down otherwise inaccessible carbohydrates, biosynthesizes essential vitamins, and alters the bioactivity and bioavailability of diverse endogenous and exogenous (xenobiotic) small molecules (13). Microbial transformations can vary significantly between individuals, affecting the toxicity and efficacy of drugs (14-18), susceptibility to infection (19, 20), host metabolism (21, 22), and disease development (Figure 1.2) (23). Nonetheless, the specific gut bacterial strains, genes, and enzymes responsible for metabolic transformations often remain poorly understood. This gap in knowledge limits the information that can be gleaned from ever-increasing amounts of microbiome sequencing data and limits our ability to manipulate or harness microbial functions to benefit human health.
Figure 1.2 Overview of gut microbial metabolism.
Gut microbes metabolize foreign, ingested molecules (xenobiotics) and endogenous compounds into metabolites with altered bioactivity and bioavailability.

Below we review current knowledge of gut microbial metabolism, focusing on its chemistry and biological consequences, and we also discuss diverse examples of known prominent microbial transformations. We primarily focus on metabolism of xenobiotics, which are ingested molecules foreign to the human body. We also discuss select examples of gut microbial metabolism of endogenous compounds, which are subject to many of the same host and microbial transformations as xenobiotics. Finally, we discuss common approaches for discovering microbial enzymes and metabolic pathways as well as strategies for manipulating gut microbial metabolism using small molecules.
1.3 Microbial and host metabolism of xenobiotics

Microbial metabolism is best understood in the context of the overall metabolic processes occurring in the human host. Orally ingested compounds pass through the upper gastrointestinal (GI) tract to the small intestine where they can be modified by digestive enzymes and absorbed by host tissues (24). Readily absorbed xenobiotics pass between or through intestinal epithelial cells, where they may be processed by host enzymes before transport to the liver via the portal vein. Following exposure to the metabolic enzymes of the liver, a main site of host metabolism in the body, xenobiotics and their metabolites enter systemic circulation, distributing into tissues and potentially affecting distal organs. In contrast, intravenously administered compounds circumvent this liver metabolism and are immediately introduced into systemic circulation. Circulating compounds are further metabolized and/or excreted, which generally occurs either via the biliary duct back into the gut lumen (biliary excretion) or through the kidneys into the urine. Metabolites returned to the gut lumen can either continue on to the large intestine, where they will eventually be excreted in the feces and potentially be modified by microbial chemistry, or they can potentially be reabsorbed by host cells in the small intestine through a process known as enterohepatic circulation.

Microbes encounter small molecules via multiple routes. Xenobiotics that are absorbed in the small intestine might avoid microbial modification altogether, although the exact biogeography of gut microbial metabolism has not been extensively investigated. In contrast, poorly absorbed xenobiotics continue through the small intestine into the large intestine and may be transformed by gut microbes. Readily absorbed compounds and compounds administered via other routes (e.g., intravenous injection) can also reach gut microbes through biliary excretion. By modifying the chemical structures of xenobiotics, gut microbes can not only influence levels of the parent
compound, but also produce metabolites with altered bioactivity and bioavailability. The products of gut microbial metabolism can be absorbed by the host and circulated systemically or interact locally with the epithelial cells lining the GI tract. Ultimately, microbial metabolites are excreted in feces or filtered by the kidneys and eliminated in the urine. In addition to directly interacting with host cells and receptors and impacting host biology, microbial metabolites can impact host enzymes involved in xenobiotic metabolism, thus influencing the fate of other compounds in the body. Thus, microbial and human metabolic processes constitute a deeply intertwined metabolic network that affects both the host and the members of the microbiota.

1.4 Contrasting the chemistry involved in host and microbial xenobiotic processing

Gut microorganisms use a broad range of enzymatic reactions to metabolize xenobiotics, many of which are distinctly microbial. In general, the gut microbiota uses hydrolytic and reductive reactions to metabolize xenobiotics (11). Key enzyme classes associated with this microbial metabolism include hydrolases, lyases, oxidoreductases, and transferases, which are among the most prevalent protein families in the gut ecosystem and are widely distributed among sequenced isolates (25, 26). While many microbial transformations may arise from relaxed substrate specificity of promiscuous microbial enzymes, recent work have revealed that some microbial enzymes may have evolved to degrade specific xenobiotics (15). In contrast to microbial metabolism, host enzymes typically use oxidative and conjugative chemistry to metabolize xenobiotics. The striking differences between host and microbial chemistry can be attributed not only to the physiological context, but also to distinct evolutionary pressures. For example, cytochrome P450s, the main liver enzymes used by the host for oxidative chemistry, may have evolved for detoxification of toxic compounds during early human evolution (27), whereas microbial metabolism could have evolved as a means of accessing alternative carbon substrates.
and electron acceptors in a densely colonized, competitive environment. The combined microbial and host metabolisms generate metabolites that would not be synthesized by the host alone and can substantially alter the bioactivities and lifetimes of xenobiotics within the human body.

1.4.1 Host enzymes involved in xenobiotic metabolism

Human xenobiotic metabolism generally transforms nonpolar compounds into hydrophilic, higher-molecular weight metabolites that are more readily excreted (Figure 1.3). This process occurs in two stages: installation or exposure of polar functional groups (“phase I”) and conjugation of these groups to more-polar metabolites (“phase II”). Phase I enzymes perform oxidative, reductive, or hydrolytic reactions to generate hydroxyl groups, epoxides, thiols, and amines. The largest class of phase I enzymes is the cytochrome P-450s, but carboxylesterases and flavin monooxygenases (FMOs) are also important in xenobiotic processing (24). Transferase enzymes predominate in phase II metabolism, appending glucuronyl, methyl, acetyl, sulfonyl, and glutathionyl groups onto xenobiotics or phase I metabolites (28). Polymorphisms in xenobiotic-metabolizing genes influence how individuals respond to both dietary and pharmaceutical interventions (29, 30).

Figure 1.3 Overview of host xenobiotic metabolism.
The oxidative and conjugative chemistry of the host increases both the size and polarity of the xenobiotic, facilitating its excretion.
1.4.2 Hydrolytic reactions

Hydrolytic chemistry is involved in the host and gut microbial breakdown of large ingested compounds, including carbohydrates and proteins, into smaller products that may be further metabolized. Hydrolase enzymes are defined by the addition of a water molecule to a substrate, followed by bond cleavage (Figure 1.4). The most abundant and relevant hydrolases in the GI tract are proteases, glycosidases, and sulfatases, and the microbiota contributes a broad range of these activities. Proteases cleave the peptide bonds linking the amino acids in polypeptide chains. The small intestine is dominated by pancreatic serine proteases from the host, whereas the colon contains many microbial cysteine- and metalloproteases (31) with different substrate specificities and potentially different clinical consequences (32). The other type of hydrolase, the glycosidases, hydrolyze glycosidic bonds using a dyad of carboxylic acid residues and a water molecule, releasing free sugars, which can be further metabolized as carbon sources (33, 34). These enzymes break down a vast diversity of glycoconjugates and oligosaccharides and are broadly distributed across gut microbes (35, 36). Single organisms can harbor hundreds of glycosidases, testifying to their important role in the gut (35). Sulfatases, which are also widespread, use the unusual amino acid formylglycine (37) to hydrolyze sulfate esters generated by phase II host metabolism and to break down sugars found on host mucus (38). Formylglycine is installed through an elaborate post-translational process by dedicated chaperones and enzymes (39). 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 (40).

In general, these widespread 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. This is seen with glucuronide conjugates of nonsteroidal anti-inflammatory drugs and the cancer therapy irinotecan (17, 41). Hydrolysis can also alter the biological activity or toxicity of xenobiotics, as observed for plant-derived glycosides like amygdalin and the artificial sweetener cyclamate (42, 43). Moreover, hydrolysis is often a prerequisite for further transformations, such as the fermentation of sugars released from indigestible polysaccharides (44), and the products of hydrolytic reactions (sugars, amino acids, and sulfate) often support microbial growth and survival in the gut.

Figure 1.4 Examples of gut microbial hydrolytic reactions. Common hydrolytic enzymes include proteases, glycosidases, and sulfatases. On the right, known enzyme substrates are shown in red and the known effects of microbial metabolism are shown in parentheses. The examples shown are selected microbial enzymes rather than an exhaustive and comprehensive list.

1.4.3 Lyases

Lyase are enzymes that break C–C or C–X bonds (where X = O, N, S, P, or halides) without relying on oxidation or the addition of water. The main lyases in the gut microbiota are polysaccharide lyases (PLs) and C–S β-lyases. PLs modify polysaccharides that contain a glycosidic bond at the β position relative to a carboxylic acid (e.g., alginate, pectin, chondroitin, and heparan) (Figure 1.5). The presence of the carboxylate carbonyl enables removal of the α
proton and subsequent β-elimination to yield an α,β-unsaturated sugar and a hemiacetal (45). A single human gut microbiome can encode for >5000 PLs (35), indicating lyase enzymes might be responsible for a vast diversity of transformations in the gut ecosystem. In contrast to PLs, 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 (Figure 1.5). These enzymes generate an aldimine linkage between the cofactor 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 ammonia and pyruvate (46). Microbes can further metabolize the thiols released in these reactions, 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 produce thiol metabolites that are further methylated and accumulate in lipophilic host tissues (47). The consequences of C–S lyase chemistry for the gut microbiota are not well understood. This activity may derive from promiscuous PLP-dependent enzymes involved in “housekeeping” functions (46); however, ammonia generated by C–S β-lyases can serve as a sole nitrogen source (48), pointing to a potential role in nutrient acquisition.
**Figure 1.5 Examples of gut microbial lyase reactions.**

The most prominent gut microbial lyases include polysaccharide lyases and C-S lyases. On the right, known enzyme substrates are shown in red and the known effects of microbial metabolism are shown in parentheses. The examples shown are selected microbial enzymes rather than an exhaustive and comprehensive list.
1.4.4 Reductive transformations

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 1.6). Reductase enzymes make use of various cofactors [e.g., NAD(P)H (i.e., NADH or NADPH), flavin, Fe–S clusters, (siro) heme, molybdenum cofactors, and other metallocofactors] to mediate the transfer of electrons or hydride equivalents (H⁺, 2e⁻) to substrates (49-51). Notably, some human gut microbes encode for among the highest number of molybdenum cofactor-dependent enzymes of any sequenced organism, suggesting these enzymes may perform particularly important roles in the gut (52, 53). Biochemical and structural characterization of gut microbial reductases have revealed that some enzymes, such as azo reductases, display broad substrate scope and reduce a specific functional group on many different molecules or can even reduce many different functional groups (54-56). Other reductases, such as the Cardiac Glycoside Reductase that reduces the cardiac drug digoxin, are specialized for distinct molecules, raising questions about the endogenous substrates of these microbial reductive transformations (15).

Reduction typically decreases the polarity of compounds and can alter charge, hybridization, and electrophilicity, which, similar to other reactions, can affect the lifetimes and activities of metabolites in the body (18, 57, 58). It is generally thought that addition of electrons to xenobiotic substrates can enable anaerobic respiration in the human gut, where oxygen is largely unavailable to serve as a terminal electron acceptor. However, experimental evidence for this hypothesis is lacking and the physiological roles of the vast majority of the reductive reactions are unclear. The observation that some prominent reductive reactions carry no obvious benefit to the metabolizing organism suggest reductive transformations may play unappreciated roles in host-microbe interactions (15). Although reductase enzymes are found in humans, many reductive
transformations are exclusively microbial (12, 18, 59) and have not yet been linked to known enzymes or organisms.

<table>
<thead>
<tr>
<th>Enzymatic reaction</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alkene reductases</strong></td>
<td>Daidzein (pathway involved in production of phytoestrogenic equol)</td>
</tr>
<tr>
<td>$\text{R}^1\text{R}^2\rightarrow \text{R}^1\text{R}^2$</td>
<td>Digoxin (decreased drug efficacy)</td>
</tr>
<tr>
<td>$2e^-; \text{NAD(P)H, FAD}$</td>
<td></td>
</tr>
</tbody>
</table>

| **Azo reductases** | Sulfasalazine (activates drug) |
| $\text{R}^1\text{N}^\equiv\text{N}\text{R}^2\rightarrow \text{R}^1^-\text{NH}_2$ | Food dyes (generates potentially toxic metabolites) |
| $4e^-; \text{NAD(P)H, FMN}$ | |

| **Sulfoxide reductases** | Sulindac (changes drug efficacy) |
| $\text{R}^1\text{SO}\text{R}^2\rightarrow \text{R}^1\text{S}^-\text{R}^2$ | |
| $2e^-; \text{Molybdenum cofactor}$ | |

| **Nitro reductases** | Nitrazepam (generates teratogenic metabolite) |
| $\text{R}^-\text{NO}_2\rightarrow \text{R}^-\text{NH}_2$ | |
| $6e^-; \text{NAD(P)H, FMN}$ | |

*Changes polarity and therefore alters bioactivity and bioavailability; reductive chemistry could be used by gut microbes in anaerobic respiration to enable growth*

**Figure 1.6 Examples of gut microbial reductive transformations.**

Gut microbial reductive chemistry involves a range of enzymes, including alkene reductases, azo-reductases, sulfoxide reductases, and nitroreductases. On the right, known enzyme substrates are shown in red and the known effects of microbial metabolism are shown in parentheses. The examples shown are selected microbial enzymes rather than an exhaustive and comprehensive list.
1.4.5 Functional group transfer reactions

Transferase enzymes use nucleophilic chemistry to move functional groups between two substrates. The gut microbiota transfers methyl and acyl groups to or from xenobiotic scaffolds (Figure 1.7). Addition reactions require chemically activated co-substrates, such as acetyl coenzyme A, adenosine triphosphate (ATP), or S-adenosylmethionine (SAM). Whereas enzymes that remove acyl groups generally rely on hydrolysis, demethylating enzymes use cofactors capable of nucleophilic catalysis [e.g., cob(I)alamin and tetrahydrofolate] (60, 61). Installation and removal of these functional groups can affect the lifetimes and bioactivities of xenobiotics in various ways. For instance, acetylation can serve as a detoxification mechanism by decreasing polarity and facilitating excretion from microbial cells, as is seen in the microbial N-acetylation of the anti-inflammatory compound 5-aminosalicylic acid (5-ASA) (62), which yields a therapeutically inactive metabolite. Host demethylation of xenobiotics liberates polar groups for further conjugation and excretion from the body (63), but in microorganisms, demethylation can provide a carbon source for growth (60).
Figure 1.7 Examples of gut microbial functional group transfer reactions.
Acetyltransferases and methyltransferases are among the most prominent gut microbial transferase enzymes. On the right, known enzyme substrates are shown in red and the known effects of microbial metabolism are shown in parentheses. The examples shown are selected microbial enzymes rather than an exhaustive and comprehensive list.

1.4.6 Radical chemistry

Radical enzymes generate high-energy intermediates containing an unpaired electron. Such processes are often oxygen-sensitive and energetically demanding, but enable microbes to perform chemically challenging reactions that are inaccessible by other modes of catalysis, including bond cleavage and formation (both C–C and C–X, where X = N, O, or halides) and skeletal rearrangements (64). Many radical enzymes used in anaerobic metabolism share a common chemical logic. By using an enzyme- or cofactor-based radical species, these enzymes typically generate a substrate-based radical intermediate through single-electron transfer or homolytic bond cleavage (Figure 1.7). This initial substrate-based radical is then converted to a product-based radical. Formation of the final product often regenerates the initial enzyme- or cofactor-based radical to complete the catalytic cycle.
Key classes of gut microbial radical enzymes include radical SAM enzymes, cobalamin (B$_{12}$)–dependent enzymes, and glycyl radical enzymes (GREs). These enzymes often mediate primary metabolism in anaerobic microbes and can directly or indirectly influence the fate of xenobiotics in the human body. Microbial GREs are widespread among individuals and represent one of the most abundant enzyme families in the human gut microbiota (25, 65, 66). Reactions catalyzed by this superfamily include generation of the disease-associated metabolite trimethylamine (TMA) by choline trimethylamine-lyase (CutC) (67), the generation of sulfite, a precursor to hydrogen sulfide by the C-S bond-cleaving enzyme isethionate sulfite-lysate (68), the decarboxylation of the tyrosine-derived metabolite $p$-hydroxyphenylacetate by $p$-hydroxyphenylacetate decarboxylase (69), and the dehydration of the host amino acid hydroxyproline by trans-4-hydroxy-L-proline dehydratase (70, 71).
Figure 1.8 Examples of gut microbial radical chemistry.
Radical enzymes perform chemically challenging reactions inaccessible by other modes of chemistry. Common reactions include bond cleavage and formation (both C–C and C–X, where X = N, O, or halides) and skeletal rearrangements. On the right, selected known substrates of radical enzymes are shown in red and the known effects of microbial metabolism are shown in parentheses.
1.4.7 Uncharacterized microbial chemistry

As highlighted above, certain microbial transformations can, with reasonable confidence, be associated with a known enzyme class. However, a range of uncharacterized metabolic activities cannot be readily explained with known biochemistry. These chemical reactions represent particularly notable opportunities for gut microbial enzyme discovery and characterization. Prominent uncharacterized transformations include racemization of the sulfone-containing drug flosequinan (72), aromatization of quinic acid derived from coffee (73), reductive cleavage of benzyl ethers in plant-derived polyphenols (74, 75), reductive cleavage of benzisoxazole heterocycles in drugs (76-78) and the removal of $p$-hydroxyl groups from catechol-containing molecules (Figure 1.9) (75, 79-82). Catechol dehydroxylation is particularly notable, for two reasons. First, this transformation, which was first reported over six decades ago (83), is not limited to a single molecule and is involved in the microbial metabolism of a range of host-derived compounds, dietary compounds, and pharmaceutical drugs (75, 81, 82, 84-86). In addition to its potential relevance to host-microbe interactions, nutrition, and pharmacology, the regiospecific removal of a hydroxyl group from the electron-rich catechol ring is an energetically challenging reaction due to the stability of the aromatic ring system. This reaction cannot be accomplished using current tools from organic synthetic chemistry. Moreover, at the start of this thesis, this reaction was not known to be catalyzed by characterized enzymes.
**Figure 1.9 Examples of uncharacterized gut microbial metabolism.**
Examples of uncharacterized microbial metabolic transformations. These reactions represent potential opportunities for enzyme discovery. On the right, selected examples are shown in red and the known effects of microbial metabolism are shown in parentheses.

### 1.5 Metabolism of dietary compounds

Diet impacts gut microbial community structure and function (87, 88), partly exerting its influence by providing a rich source of molecules that are subject to microbial metabolism and impact microbial growth (89, 90). The types and extent of microbial modifications of dietary compounds vary substantially among individuals, presumably due to differences in the presence and abundance of gut microbial enzymes, and the bioactivities of the resulting metabolites range from beneficial to acutely toxic. As much attention has been focused on microbial metabolism of complex, plant-derived polysaccharides (34, 91, 92), we have chosen to highlight transformations of noncarbohydrate dietary components.
1.5.1 Dietary protein

Dietary protein is necessary for supplying humans with essential amino acids, but the source and amount of protein vary depending on the diet. The gut lumen is rich in both host and microbial proteases, and studies increasingly indicate that differential microbial proteolytic activity may directly contribute to human disease. For example, the gut microbiota is associated with celiac disease (CD), a common autoimmune disorder characterized by an inflammatory response to dietary gluten found in wheat-based foods (93). This proline-rich protein evades complete digestion by host proteases, resulting in the generation of high-molecular weight, immunogenic peptides. The gut microbiota may affect CD by altering gluten proteolysis. Fecal suspensions from healthy individuals and CD patients process gluten proteins and immunogenic peptides differently (94). Moreover, gluten-derived peptides generated by Pseudomonas aeruginosa, an opportunistic pathogen in CD patients, are prone to translocation across the mouse intestine and elicit an enhanced gluten-specific immune response in comparison with peptides produced by Lactobacillus spp. from healthy individuals (95).

Gut microbes can also metabolize amino acids obtained from dietary protein, including L-phenylalanine, L-tyrosine, and L-tryptophan, into a range of bioactive products (Figure 1.10) (96-99). For example, gut bacteria metabolize L-tryptophan into many products, including the antioxidant indole-3-propionic acid, the neurotransmitter tryptamine, and indole, the latter of which can undergo hydroxylation and sulfation by hepatic enzymes to generate the uremic toxin indoxyl sulfate (97, 100, 101). In addition, gut microbes metabolize L-tyrosine into the bioactive amine p-tyramine, which promotes host serotonin biosynthesis (102), and into p-cresol, which, upon sulfation by the liver, influences clearance of the drug acetaminophen (Figure 1.10) (103).
Figure 1.10 Gut microbial metabolism of amino acids derived from dietary protein.
The gut microbiota can break down larger proteins into free amino acids, which are further metabolized. A) Microbial metabolism of tyrosine produces \( p \)-tyramine, a biogenic amine that promotes host serotonin biosynthesis. A different tyrosine metabolic pathway, which involves reductive deamination, oxidation, and subsequent decarboxylation, produces \( p \)-cresol, a metabolite that interferes with host xenobiotic metabolism upon its sulfation in the liver. B) Microbial decarboxylation of L-tryptophan produces tryptamine, a neurotransmitter that influences gut motility. A separate pathway produces indole from dietary L-tryptophan. Following oxidation and sulfation by the liver, indole becomes indoxyl sulfate, a uremic toxin. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.

1.5.2 Dietary lipids

Variable gut microbial metabolism of lipids and lipid-derived compounds is associated with a variety of human diseases \((104, 105)\). One notable example involves dietary cholesterol, a major component of Western diets that is associated with increased risk of cardiovascular disease \((106)\). While ingested cholesterol is absorbed in the small intestine and subsequently undergoes biliary excretion and enterohepatic circulation, gut microbial reduction of cholesterol generates coprostanol, which cannot be reabsorbed and is excreted. This transformation may reduce levels of circulating cholesterol. Coprostanol comprises up to 50% of the steroids in human
feces (107), and GF mice colonized with microbes from high- and low-cholesterol–reducing patients produce distinct amounts of coprostanol (108). Studies of the cholesterol-reducing gut bacterium Eubacterium coprostanoligenes isolated from a pig sewage lagoon indicate that coprostanol synthesis may involve oxidation to 5-cholest-3-one followed by alkene isomerization to 4-cholest-3-one, conjugate reduction, and ketone reduction (109). Another prominent example of lipid metabolism involves dietary phosphatidylcholine. Diverse organisms harboring phospholipase D enzymes hydrolyze this lipid and release choline, which can then be further metabolized into the disease-associated metabolite TMA (Figure 1.11) (110). Phospholipase D enzymes from human gut bacteria differ from those found in pathogens, which are involved in penetration of host cells and virulence, suggesting that the gut bacterial enzyme might have evolved specifically for lipid metabolism in a commensal context (110).

Figure 1.11 Gut microbial metabolism of dietary lipids.
A) Microbial metabolism reduces cholesterol into coprostanol, but the precise steps in the pathway are not known. The pathway may involve direct reduction or a multi-step oxidation-reduction pathway. B) Microbes can also hydrolytically release choline from dietary phosphatidylcholine. Free choline is a nutrient for both host and microbes. Further metabolism by host and microbes converts this choline into trimethylamine N-oxide (TMAO), a metabolite associated with a number of diseases. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.5.3 Dietary phytochemicals

In addition to contributing macronutrients such as proteins and carbohydrates, the human diet is a rich source of bioactive small molecules from plants. These dietary phytochemicals have been linked to the health benefits of consuming certain plants (111). Many phytochemicals are poorly absorbed by the host and instead reach the large intestine where they encounter bacteria, suggesting that their biological effects may derive from interactions with the gut microbiota (Figure 1.12) (112). Dietary phytochemicals are extensively metabolized by the gut microbiota, and this metabolism can vary widely between individuals (113). While some transformations have been linked to individual bacterial strains, many reactions remain understood only at the level of the microbial community. Moreover, the vast majority of enzymes responsible for dietary phytochemical metabolism are unknown. Linking known transformations to individual genes and enzymes could elucidate whether microbial phytochemical metabolism can directly affect host biology or are biomarkers for disease susceptibility.

In many cases, microbial metabolism is required for converting ingested compounds into bioactive metabolites. This is illustrated in the microbial metabolism of glucosinolates, a class of molecules found in Brassica vegetables such as cabbage, kale, and broccoli. Brassica plants biosynthesize glucosinolates as defense chemicals to fend off potential predators. Upon tissue damage, the plant enzyme myrosinase hydrolytically cleaves glucosinolates to generate isothiocyanates (ITCs), which are potent sulfur-and nitrogen containing small molecules with a number of established biological effects (114). ITCs inhibit host CYP450s, induce of expression of glutathione-S-transferases, display anti-tumor activity in vitro, and interfere with pathways involved in pain signaling and inflammation (115, 116). Importantly, however, the bioactivation of glucosinolates does not just occur in damaged plants, but also inside the human body due to the
action of gut microbial enzymes (Figure 1.12) (117, 118). Multiple human gut microbes hydrolyze glucosinolates into ITCs (89), and recent work by Sattely and co-workers identified and characterized Bacteroides enzymes involved in this bioactivation (119). Although microbial glucosinolate hydrolysis results from promiscuous enzyme activity, experiments colonizing gnotobiotic mice with either wild-type or Bacteroides fragilis strains indicated that these enzymes are sufficient for bioactivation in vivo. The identification of gut microbial ITC-producing enzymes will enable future studies exploring the link between this activity and host health and could uncover whether microbial metabolism may account for links between a Brassica-rich diet and human health (120).

Like glucosinolates, phenolic compounds constitute an important class of phytochemicals that are metabolized by human gut microbes and are associated with host health. Commonly consumed phenolic compounds include caffeic acid, a universal lignin precursor in plants, soy isoflavones, lignans from flaxseed and sesame seeds, flavonoids like the catechins and gallocate esters found in tea and chocolate, and ellagic acid from nuts and berries (85, 111, 113, 121). Phenolic compounds are poorly absorbed and reach the lower intestine, where microbial metabolism subjects these molecules to ring cleavage, demethylation, and dehydroxylation (122, 123), producing metabolites with higher bioavailability, increased bioactivity, and a correlation with lowered disease risk (Figure 1.12) (111, 113). For example, gut microbial metabolism of the polyphenol ellagic acid, a pathway that involves dehydroxylation of a catechol motif, produces urolithins that promote gut barrier function, prolong lifespan in C. elegans, and increase muscle function in rodents (124, 125). Similarly, the multi-step metabolism of flaxseed lignans into the phytoestrogens enterodiol and enterolactone, a pathway that also involves catechol dehydroxylation, has been linked to decreased tumor burden in mice (74, 85). Finally,
Dehydroxylation of catechin and quercetin produces metabolites that can directly stimulate the immune system and protect against disease (126, 127).

Although microbial phenol metabolism is generally considered beneficial, the microbial metabolites produced can also interfere with host processes. For example, deglycosylated polyphenols inhibit the host enzyme catechol-O-methyltransferase in vitro and also impede clearance of catecholic drugs such as L-dopa (128, 129). Naringin, a flavanone present in grapefruit is metabolized by human fecal samples to the aglycone naringenin, a potent inhibitor of CYP3A4 (130). Grapefruit juice is known to alter the pharmacokinetics of over 20 common drugs, and although the precise contributions of microbial metabolism to naringin hydrolysis in vivo remains undefined, this finding raises the possibility that differing levels of naringenin-producing gut microbes could explain the variable effects of grapefruit juice on human drug metabolism (131). Similarly, the flavonoid ginsenoside found in ginseng has low bioactivity, whereas the microbially deglycosylated compound potently inhibits CYP450s in vitro and interferes with host xenobiotic clearance (132). While it is known that bacteria de-glycosylate dietary polyphenols to access sugars for growth (133), specific gut microbial enzymes involved in this metabolism are generally unknown.
Figure 1.12 Gut microbial metabolism of dietary phytochemicals. Microbial metabolism of plant-derived compounds produces metabolites with altered bioactivity.
Figure 1.12 (Continued)
A) Microbial hydrolysis of glucosinolates present in in Brassica plants produces bioactive isothiocyanates. B) Involving benzyl ether reduction, demethylation, and catechol dehydroxylation reactions, the multistep microbial metabolism of dietary lignans produces phytoestrogens such as enterodiol. These metabolites are linked to lower risk of hormone-dependent cancers. C) The multi-step metabolism of dietary ellagic acid present in nuts and berries involves initial hydrolysis followed by decarboxylation and diverse dehydroxylation reactions. The ellagic acid metabolites urolithins have a number of bioactivities, ranging from promoting gut barrier function to improving muscle function in mice. D) Dietary caffeic acid, a universal lignin precursor in plants, is first reduced and then dehydroxylated into a metabolite with unknown biological effects. This pathway, as well as those in A) and B), involves catechol dehydroxylation, a prominent gut microbial transformation. E) Finally, gut microbial hydrolysis of the grapefruit-derived polyphenol naringin produces naringenin, a potent inhibitor of host CYPs that could potentially account for grapefruit’s inhibitory effects on host drug clearance. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.

1.5.4 Molecules produced during cooking

An understanding of gut microbial metabolic activities could provide new insights into the biological consequences of cooking practices and food preparation. Cooking has a major impact on the gut microbiota by increasing starch accessibility and changing the pool bioactive small molecules present in the food (134). Microbes metabolize molecules produced during cooking and can change their biological lifetimes and activities (Figure 1.13). For instance, the mutagenic potential of heterocyclic amines, poorly absorbed molecules produced during charring of meat and fish, can be altered by gut microbial metabolism (135). Gut microbes convert 2-amino-3-methylimidazo[4,5-f]quinoline (IQ) into the potential mutagen 7-hydroxy IQ (136) and can hydrolyze IQ-glucuronide conjugates, prolonging the lifetime of IQ in the body (137). Gnotobiotic rats mono-associated with an Escherichia coli strain encoding a β-glucuronidase (uidA) have higher levels of unconjugated IQ and increased colonic DNA damage compared with rats colonized with an isogenic uidA mutant (Figure 1.13) (137). These findings implicate the gut microbiota in the known link between charred meat and cancer (138). Gut microbes also
metabolize fructoselysine, a Maillard reaction product (MRP) derived from reaction between amino acids and reducing sugars during high-heat cooking (139). Collinsella species transcriptionally induce the metabolizing genes in response to this molecule, indicating that gut microbes may have specifically evolved to metabolize this MRP (Figure 1.13). The metabolism provides a carbon source to Collinsella intestinalis, increasing its abundance in mice fed a highly processed diet (139). This suggests that food processing can directly impact microbial fitness and highlights the importance of considering not just the ingredient itself, but also how it is prepared, to understand diet-microbe interactions.

Figure 1.13 Gut microbial metabolism of molecules produced during cooking. Microbes metabolize dietary molecules produced during cooking. A) The Maillard reaction product fructoselysine can be shunted into primary metabolism, generating glucose-6-phosphate which provides a carbon source to specific gut microbes B) Gut microbial hydrolysis of glucuronidated 7-hydroxy-IQ releases the potent mutagen 7-hydroxy-IQ in the gut. Known effects of microbial metabolism are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.5.5 Food additives

In addition to metabolism of naturally occurring molecules, human gut microbes interact with components of our diets that are added in the process of food manufacturing (e.g., artificial sweeteners, emulsifiers, and preservatives). For instance, although many artificial sweeteners are poorly metabolized by humans, studies demonstrate that they are susceptible to microbial transformation. Gut microbes convert the artificial sweetener cyclamate into cyclohexylamine via hydrolytic cleavage of its sulfamate linkage. Cyclamate was banned in the United States after studies suggested that cyclohexylamine was carcinogenic, and both this finding and the continued use of this sweetener remain controversial (140). A cyclamate hydrolyzing enzyme has been partially purified from a guinea pig–associated strain, but human gut microbial hydrolases with this activity have not been identified (43). Gut microbes can also metabolize the artificial sweeteners stevioside and xylitol using unknown enzymes (141, 142). The gut microbiota acquires the ability to transform xylitol and cyclohexamate after prolonged exposure, suggesting that long-term ingestion of dietary components can select for particular microbial metabolic functions (143). Consistent with this idea, certain strains of the intestinal pathogen Clostridium difficile have acquired the ability to grow on low concentrations of trehalose, a sugar substitute and common additive introduced in the 2000s (144). Strikingly, the emergence of these strains correlates with the introduction of trehalose in the geographic location where the strains were isolated (144).

Often used as coloring agents in processed foods, azo dyes represent an additional group of common food additives. Gut microbes reductively cleave the azo linkages in these dyes, generating aniline products (Figure 1.14). Despite their use for >150 years as textile dyes, food colorings, and pharmaceuticals, we have an incomplete understanding of the organisms and enzymes that process these molecules. Azo reduction is performed by azoreductases, promiscuous
flavin- or NAD(P)H-dependent enzymes found in many eukaryotes and bacteria, including *P. aeruginosa*, *Bacillus subtilis*, *Enterococcus faecalis*, and *Staphylococcus haemolyticus* (145). Reduction of seven FDA-approved azo dyes by *Clostridium* species did not generate mutagenic metabolites *in vitro* (49) and no toxicity was observed in mice fed high daily doses of these compounds (146). However, the specific effects may depend on the individual dye backbone (147). Over >3000 derivatives currently in use (148) and understanding how these are transformed by gut microbes could unveil new links between processed foods and human health.

*Figure 1.14 Gut microbial metabolism of food additives.*

Microbial metabolism of food additives can change their gastronomic properties and biological activities, affecting human health. A) The commonly used artificial sweetener cyclohexylamine is hydrolyzed into a metabolite with unknown biological effects. B) Commonly used azo food dyes undergo microbial reduction. Further host metabolism of the microbially generated aniline products generates potentially mutagenic metabolites. A common example involves the yellow dye tartrazine. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.6 Metabolism of pharmaceuticals

The human gut microbiota is known to transform hundreds of pharmaceuticals, spanning many indications and host targets, into metabolites with altered pharmacological properties \((11, 149, 150)\). In some cases, the teratogenic \((151)\), toxic \((17, 41)\), and even lethal \((152)\) effects of these microbial metabolic activities were not recognized until drugs were on the market. Ongoing studies have illuminated a complex interplay between drugs and gut microbes, including effects of drugs on microbial growth and effects of microbes on immunotherapy efficacy \((16, 153-155)\), but we focus here on examples of direct microbial modifications of drugs.

1.6.1 Anti-inflammatory and GI agents

Multiple drugs that target the GI tract are affected by gut microbes, either by direct chemical modification or indirectly through the many interactions these organisms have with host cells in this environment. Notably, several of these agents rely on microbial metabolism for converting inactive precursors (prodrugs) to pharmaceutically active compounds (Figure 1.15). Prominent examples are anti-inflammatory drugs that contain azo linkages, including the inflammatory bowel disease (IBD) medication sulfasalazine \((57)\). Gut microbes reduce sulfasalazine into sulfapyridine and the active anti-inflammatory agent 5-ASA, and various intestinal bacteria can further metabolize 5-ASA into N-acetyl 5-ASA, a metabolite that lacks anti-inflammatory activity (Figure 1.15). Considerable variability in acetylation rates has been observed in human fecal samples \((156)\). Together with differences in azo reduction, this observation could potentially explain variable therapeutic efficacy of sulfasalazine in patients. N-acetyl ASA inhibits the growth of anaerobes, including Clostridium difficile \((62)\), which suggests that this activity could affect gut microbiota composition. However, the target of 5-ASA has not been clearly identified. Other gut microbial activities involved prodrug activation include
reduction of the sulfoxide found in the anti-inflammatory compound sulindac (157) and reduction of the N-oxide of the anti-diarrheal drug loperamide (Figure 1.15) (58). Gaining a better understanding of the specific organisms and enzymes responsible for these activities and their presence in patients could aid in drug selection and dosing.

Figure 1.15 Gut microbial metabolism of anti-inflammatory and GI agents. A) The microbial reduction of the azo bond of the anti-inflammatory drug Sulfasalazine generates 5-aminosalicylic acid (5-ASA), the active agent that decreases inflammation in IBD. It also generates the antibiotic Sulfapyridine. Gut microbes can further acetylate 5-ASA, producing a metabolite that no longer is active in suppressing inflammation but can inhibit growth of the intestinal pathogen *C. difficile*. B) Similar to sulfasalazine, loperamide-N-oxide requires microbial action to be activated in the gut. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.6.2 Immunosuppressants

Gut microbes also metabolize immunosuppressants, including tacrolimus, a drug that is commonly prescribed to kidney transplant patients. Taken orally, tacrolimus has a narrow therapeutic window and also displays remarkable interindividual variability in bioavailability, ranging from 5-93% in patients (158). Recent observations have associated gut microbial community composition with patient dosage, suggesting a potential role of the gut microbiota in drug efficacy (159). Recently, Jeong and co-workers discovered that many gut bacteria can directly degrade tacrolimus through ketone reduction, producing a metabolite that displays 15-fold lower immunosuppressant activity compared to tacrolimus (Figure 1.16) (160). Although this route of metabolism had not been reported previously, the authors detected the microbial metabolite in stool samples from kidney transplant patients, revealing that this previously unknown metabolism indeed occurs in vivo. However, the clinical relevance of this metabolism and the gut microbial enzymes responsible remain unknown.

While microbial tacrolimus metabolism was discovered long after the drug had reached the market, other immunomodulating drugs were already known to undergo microbial metabolism at the preclinical stage. Fostamatinib, a spleen tyrosine kinase inhibitor, was known to undergo gut microbial metabolism in 2010, prior to its approval in 2018 for the treatment of immune thrombocytopenic purpura (81). This molecule is subject to both host and microbial processes, and mass balance studies in humans indicate that the microbial transformation involves initial demethylations of two aromatic methoxy groups, followed by dehydroxylation of a pyrogallol motif, producing a 3,5-diol metabolite that is found in fecal incubations and in human subjects consuming the drug (Figure 1.16) (81). Although the precise biological consequences of the microbial metabolism are unknown, it is notable that this drug undergoes aromatic
dehydroxylation, a prominent transformation that is also involved in the microbial metabolism of host neurotransmitters and ingested dietary plant-derived compounds. However, the organisms and enzymes responsible are unknown. Eltrombopag, a drug that, like Fostamatinib, is used to treat immune thrombocytopenic purpura, is also metabolized by the gut microbiota (161), highlighting that gut microbes can transform different medications used to treat a single condition.

![Figure 1.16 Gut microbial metabolism of immunosuppressants.](image)

A) Microbial metabolism of tacrolimus, an immunosuppressant prescribed to patients undergoing kidney transplant, produces a metabolite with 15-fold lower potency. B) Similarly, host-and microbial co-metabolism of the spleen tyrosine kinase inhibitor Fostamatinib (used to treat autoimmune disease) produces a metabolite with unknown activity. Fostamatinib metabolism involves microbial aromatic dehydroxylation, a prominent reaction that also features in the microbial metabolism of neurotransmitters and diet-derived phenols. This metabolism was discovered before the drug was approved for used in humans in 2018. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.6.3 Cancer chemotherapy

Patient response to cancer treatment can differ dramatically between individuals, both in terms of efficacy as well as the severity of side effects, and emerging studies suggest that differences in the gut microbiota may contribute to this phenomenon. In addition to modulating patient responses to immunotherapy drugs such as checkpoint inhibitors, microbes can directly alter the structures of cancer chemotherapy drugs and their metabolites, affecting their interactions with host cells. Microbes have a high potential for modifying the chemical structures of commonly used chemotherapeutics (Figure 1.17) (162). For example, co-incubation with *E. coli* or *Listeria welshimeri* either increased or decreased the efficacy of half of a panel of 30 anticancer drugs toward cancer cell lines. Assays with *E. coli* and a subset of these drugs (gemcitabine, fludarabine, and CB1954) revealed evidence for direct chemical modification by the bacteria. Following up on these observations, more recent work has revealed a potential role for bacteria in mediating tumor resistance to gemcitabine. Bacteria in the tumor microenvironment metabolize this drug into the inactive metabolite 2′,2′-difluorodeoxyuridine using the long isoform of the enzyme cytidine deaminase (Figure 1.17) (163). This microbial metabolism contributed to gemcitabine resistance in a mouse model of colon carcinoma and the finding that Gammaproteobacteria were detected in a majority of patient tumor samples implicates microbial metabolism in potential drug resistance in humans.

In addition to its impact on drug efficacy, microbial metabolism of cancer therapeutics also influences side effects and toxicity. A prominent example involves the colon cancer drug irinotecan. Host carboxylesterases transforms this intravenously administered prodrug into a metabolite named SN-38, the active, neoplastic topoisomerase I poison. As part of host liver metabolism, SN-38 can be further glucuronidated by host enzymes into the inactive metabolite
SN-38G, which is excreted into the GI tract. In the gut, however, microbes can remove the glucuronide group, regenerating the toxic chemotherapeutic SN-38 which causes dose-limiting diarrhea in patients (17). This metabolism will be discussed in further detail in section 1.9.1 below. Similarly, gut microbes metabolize doxorubicin, a chemotherapeutic associated with severe side effects such as cardiac toxicity and damage to the gut epithelium. Gut microbial metabolism may be beneficial in this context, as the reductive glycosylation and further reduction of doxorubicin produces metabolites that decrease drug toxicity and improve survival of the model organism *C. elegans* (Figure 1.17) (164). Doxorubicin metabolism in *E. coli* and the related proteobacterium *Raoutella planticola* depends on the cofactor molybdopterin, but the specific enzymes responsible for this reaction remain unknown. Finally, unknown gut organisms and enzymes metabolize the chemotherapeutic agent methotrexate into a less toxic metabolite (165). Co-administering methotrexate with antibiotics was lethal in animals (166) and patients (167), suggesting a role for microbes in drug toxicity. Altogether, these examples underscore the pivotal role of gut microbial metabolism in chemotherapeutic toxicity and side effects.
Figure 1.17 Gut microbial metabolism of cancer chemotherapy drugs.

A) The cancer drug irinotecan is metabolized by the host into the active antineoplastic agent SN-38. This can be further glucuronidated by host glucuronide transferases and excreted into the gut, where gut microbial β-glucuronidases reactivate the drug and contribute to dose-limiting diarrhea.

B) Microbial metabolism of the chemotherapeutic Gemcitabine produces an inactive metabolite. This metabolic pathway may contribute to tumor drug resistance in patients.

C) The two-step microbial metabolism of Doxorubicin decreases drug toxicity, potentially limiting the severe side effects associated with this chemotherapeutic. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.6.4 Cardiovascular drugs

Gut microbes also metabolize drugs prescribed to prevent or treat cardiovascular disease. Commonly consumed for cardiovascular disease prevention, statins lower cholesterol levels by inhibiting 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase, the rate-limiting enzyme in cholesterol synthesis. There is considerable interindividual the the ability of statins to lower LDL-cholesterol in human patients, with more than 50% of patients experiencing a suboptimal response (<30% reduction in LDL-cholesterol) to a 10 mg dose of the commonly prescribed simvastatin (168). Emerging evidence suggests that the gut microbiota could potentially contribute to this phenomenon. For example, several microbial metabolites, including the reduced cholesterol product coprostanol and secondary bile acids, correlate with response to simvastatin in human patients (169). Furthermore, human intestinal microbes directly metabolize simvastatin in ex vivo incubations. The proposed pathway involves de-esterification to release dimethylbutanoic acid, which can be further metabolized through hydroxylation and demethylation (Figure 1.18). Alternatively, simvastatin can be subjected to C-C and C-O bond cleavage, producing 2-hydroxyheptanoic acid, which is subject to further metabolism (Figure 1.18) (170). The biological effects of the statin metabolites and the relevance of the microbial metabolism in vivo are unknown. Notably, statins also possess antimicrobial activity and may alter gut microbiota composition and function, potentially affecting drug metabolism and patient response (171). In addition, the common gut bacterium Eggerthella lenta reductively inactivates the cardiac drug digoxin, reducing efficacy (Figure 1.18) (172). The discovery of the genes and enzymes involved in digoxin metabolism as well as their clinical implication will be further discussed in section 1.8.2.3 of this chapter.
Figure 1.18 Gut microbial metabolism of cardiovascular drugs.
A) The cholesterol-lowering drug simvastatin is metabolized via multiple routes. B) The prominent human gut Acitnobacterium Eggerthella lenta uses a Cardiac Glycoside Reductase to transform the cardiac drug digoxin into an inactive metabolite. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.6.5 Anti-viral drugs

Microbial metabolism of anti-viral compounds can cause deleterious drug-drug interactions in the host. Such interactions have been observed in the coadministration of the anti-viral drug sorivudine with the chemotherapeutic agent 5-fluorouracil (5-FU) (152). Gut microbes hydrolyze the glycosidic linkage of ingested sorivudine to release \((E)-5-(2\text{-bromovinyl})\text{-uracil (BVU)},\) which is absorbed into circulation (Figure 1.19). The host enzyme dihydropyrimidine dehydrogenase (DPD) then further reduces BVU, generating an allyl bromide intermediate that is attacked by an active cysteine of DPD to form a covalent thioether adduct (173). Because DPD participates in 5-FU degradation, its irreversible inhibition by BVU leads to 5-FU accumulation and toxicity (174). This microbiota-dependent drug-drug interaction is implicated in the deaths of 18 patient receiving soruvidine and 5-FU (152, 174). BVU is also generated from the hydrolytic metabolism of the anti-viral drug Brivudine. Brivudine metabolism involves host and microbial processes, and recent work from Goodman and co-workers has linked this metabolism to a predicted nucleoside phosphorylase in Bacteroides species (14) (Figure 1.19). Gnotobiotic mice colonized by \(B.\) thetaiotaomicron wild-type and mutant strains differing only in the presence of this enzyme have different pharmacokinetic profiles, and pharmacokinetic modeling suggested that the systematic exposure to the toxic BVU metabolite varies substantially depending on bacterial enzymatic activity (14). These results highlight that microbial enzymes should be considered in drug-drug interactions.
Figure 1.19 Gut microbial metabolism of anti-viral drugs.
The anti-viral drugs Brivudine and Sorivudine can both be hydrolyzed into bromovinyluracil (BVU), a metabolite that implicated in a lethal drug-drug interaction with the chemotherapeutic 5-fluorouracil. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.

1.6.6 Central Nervous System (CNS) drugs

Many CNS-targeted drugs are administered orally, and gut microbial metabolism affects absorption and circulation of active drugs and metabolites (Figure 1.20). For example, microbiota-mediated reductive cleavage of an isoxazole heterocycle inactivates the anti-convulsant zonisamide (78). Nitro group reduction by gut microbes and acetylation by host enzymes converts the anti-anxiety medication nitrazepam (NZ) into the teratogen 7-acetylanitrazepam (Figure 1.19) (175). This compound induced fetal abnormalities in rats (151), and very high exposure to NZ during pregnancy increased rates of congenital disease in children, suggesting that microbial
metabolism of CNS-targeted drugs could have potentially deleterious effects (176). Other notable examples of microbial CNS-targeted drug metabolism involve methamphetamine (177) and the anti-depressant risperidone (76). The majority of enzymes involved in CNS-targeted drug metabolism are obscure. As studies continue to reveal connections between the gut microbiota and various neurological diseases (178), as well as between psychopharmacological drugs and microbial growth and community structure (153, 179) it will become increasingly important to characterize microbial interactions with CNS-targeted drugs at the molecular level.

While some drugs are subject to single microbial modifications, others participate in metabolism that involves multiple steps. A prominent example of multi-step microbial drug metabolism involves L-dopa. The metabolism of this drug will be discussed in more detail in chapter 2, section 2.1. Introduced in the 1960s, 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 the host enzyme aromatic amino acid decarboxylase to restore dopamine levels (180). However, extensive metabolism within the gut by both host and microbial enzymes affects the concentration of drug reaching the brain because dopamine generated in the periphery can no longer cross the blood-brain barrier (Figure 1.20). Microbial metabolism was first reported in 1957 in rabbits and rats, years before the drug had reached the market (83). The main microbial pathway, which was estimated to account for 5% of total metabolism based on metabolite profiling in rats (181), involves decarboxylation and p-dehydroxylation to convert L-dopa to m-tyramine, which can be further oxidized to m-hydroxyphenylacetic acid by the host (Figure 1.19) (79, 84, 182-186). Differences in these activities may contribute to the substantial variation observed in patient response to L-dopa (187). Although a tyrosine decarboxylase from a food-associated strain of Lactobacillus
brevis accepts L-dopa in vitro (188), the human gut microbial enzymes responsible for L-dopa metabolism were unknown at the start of the work described in this thesis. Identifying these enzymes could not only add to our knowledge of how CNS-targeted drugs are metabolized in the body but could also provide potential new biomarkers and therapeutic targets for increasing L-dopa efficacy.

Figure 1.20 Gut microbial metabolism of Central Nervous System (CNS) drugs. A) The reductive cleavage of an isoxazole heterocycle inactivates the commonly prescribed anti-convulsant Zonisamide. B) Microbial and host co-metabolism transforms the anti-anxiety drug Nitrazepam into 7-aminoacetylnitrazepam, a teratogen that causes birth defects in animal models.
Figure 1.20 (Continued)
C) Microbes metabolize L-dopa, the main drug treatment for the neurodegenerative condition Parkinson’s disease. Taken orally, this drug crosses the blood-brain barrier and is converted by the host enzyme amino acid decarboxylase into dopamine, the active therapeutic agent. However, a significant amount of L-dopa is metabolized in the gastrointestinal tract. Both host and microbial metabolic activities decarboxylate L-dopa into dopamine. This neurotransmitter cannot cross the blood-brain barrier and instead causes side effects in the periphery. Microbial dopamine dehydroxylation produces m-tyramine, a metabolite that has unknown biological effects and is excreted in urine of patients. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.

1.6.7 Biologics and peptide drugs

Although most biologics are delivered intravenously, smaller peptides and insulin (189) are available as oral formulations. Biliary excretion has also been observed for short peptides (190). Gut microbes from rats degrade the peptide drugs insulin, calcitonin, and azetirelin (32, 191), and addition of protease inhibitors or antibiotics to caecal suspensions reduced metabolism. A panel of peptides also showed variable levels of degradation by human fecal samples (31). Peptide-based therapeutics and biologics are growing in popularity, which means that elucidating their interactions with gut microbes will become increasingly relevant.

1.6.8 Herbal supplements and traditional medicines

Finally, gut microbes transform poorly absorbed constituents of herbal and traditional remedies, which are sometimes used in place of approved drugs to prevent or treat disease (Figure 1.21). Microbial metabolism of these traditional remedies can have deleterious consequences. Amygdalin is a mandelonitrile glycoside found in almonds and fruit pits that was used as an alternative cancer treatment in the 1960s (192). Clinical trials demonstrated no improvements in cancer survival or symptoms (193), and many patients instead suffered from acute cyanide poisoning (194). Animal experiments showed that gut microbes hydrolyze the glycosidic linkage of amygdalin to release mandelonitrile, which spontaneously breaks down to produce
benzaldehyde and cyanide that causes toxicity (42, 195). Gut microbes also metabolize berberine, a plant-derived benzoisoquinoline alkaloid that has been traditionally used in Chinese medicine to treat gastrointestinal disorders. In contrast to the toxicity-associated metabolism of amygdalin, gut microbes metabolize berberine in a way that allows intestinal absorption (196, 197). Because herbal supplements and traditional medicines are not regulated to the same extent as pharmaceuticals, their modes of action are typically less well characterized. Elucidating how gut microbes process these compounds may contribute extensively to our fundamental understanding of their effects.

**Figure 1.21 Gut microbial metabolism of herbal supplements.**

Gut microbes transform poorly absorbed constituents of herbal and traditional remedies, which are sometimes used in place of approved drugs to prevent or treat disease. A) A prominent example involves amygdalin, a molecule present in fruit pits that was used as an alternative cancer treatment. Microbial hydrolysis produces a mandelonitrile, a metabolite that spontaneously decomposes into cyanide, causing side effects and even death in patients. B) Microbes also metabolize the herbal supplement berberine into dihydroberberine, a metabolite with significantly higher bioavailability. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.7 Metabolism of host-derived compounds

As discussed above, ingested compounds represent a major group of molecules that are metabolized by gut microbes. However, the gut microbiota also makes substantial contributions to the metabolism of endogenous molecules produced in the human body, both by regulating host metabolism and by directly modifying host-derived compounds (13). The host and microbial chemistry involved in the transformations of endogenously produced molecules is similar to the chemistry involved in xenobiotic metabolism, and the consequences range from decreasing levels of the parent compound to production of metabolites with altered bioactivity and availability (13). Below we highlight some prominent examples of direct microbial modifications of host-derived compounds.

1.7.1 Mucus components

The host epithelium in the gut is covered by mucus, a hydrogel mainly comprised of heavily glycosylated proteins (mucins) and lipids. Mucins are secreted by specialized Goblet cells in the gastrointestinal tract (198), with MUC2 being the major mucin in the human intestine (199). In addition to serving as a physical barrier that protects the host from microbial infringement and decreases the virulence of potential pathogens, the mucus layer provides a rich source of metabolic substrates for gut bacteria (Figure 1.22) (200-202). Gut Bacteroides species encode many hydrolytic enzymes specialized for liberating sugars found in the glycosylated mucins, including sialic acid, heparan sulfate, and fucose (Figure 1.22) (34, 35, 92). Once released from mucus, these carbohydrates are made available to metabolizing bacteria and to the community as a whole, affecting community composition (203). Microbial metabolism of mucin glycoproteins can have beneficial effects but can also weaken the mucus barrier and lead to inflammation (204, 205). Enzymes involved in this sugar metabolism include sulfatases and glycosidases, and many of these
enzymes have been biochemically characterized (206-208). Gut microbes also degrade phosphatidylcholine, a main lipid component of mucus. Hydrolysis of this molecule by microbial PLD enzymes releases choline, a carbon source that can be further transformed into the disease-associated metabolite TMA (110). This metabolism may have the potential to weaken the mucus barrier, an observation that is particularly noteworthy given the increasingly appreciated interaction between the gut microbiota, mucus barrier, and inflammation in IBD pathogenesis (209). Because the extent of mucus degradation varies based on host diet, dietary interventions may provide opportunities to carefully control the beneficial and deleterious consequences of microbial mucus metabolism (210).
Figure 1.22 Gut microbial metabolism of mucus components.
Host mucus is mainly comprised of heavily glycosylated proteins (mucins) secreted by specialized Goblet cells into the gastrointestinal tract. MUC2 is the main mucin protein that makes up the intestinal mucus barrier. Glycosylation starts at serine or threonine residues and extends into long chains comprised of diverse sugars. Microbial sulfatases, fucosidases, sialidases are prominent enzyme classes involved in the metabolism of these sugars (the bonds that these enzymes cleave are indicated in red, italic text). Released sugars serve as carbon sources and are available to metabolizing bacteria or the microbial community as a whole. The glycosylation pattern shown above represents a schematic overview. There is a vast diversity of different glycosylation patterns possible in practice.
1.7.2 Bile acids

Microbes also metabolize host molecules involved in digestion. A well-characterized interaction between host digestive fluids and microbes involves bile, a green fluid synthesized in the host liver and secreted by the gall bladder (211, 212). Bile mainly consists of cholesterol and cholesterol derivatives called bile salts (213). Bile salts are highly amphipathic, a property that enables these molecules to act as detergents and solubilize otherwise insoluble fats and vitamins (214) and to serve as antimicrobials (215). Found at millimolar concentrations in the gut lumen, bile salts directly interact with the gut microbiota and are subject to microbial metabolism (Figure 1.23) (216). In the first step of microbial metabolism, bile salt hydrolase enzymes hydrolytically cleave amide bonds that link conjugated amino acids (usually taurine or glycine) to the bile acid core (21). The mechanism involves nucleophilic attack by an active site cysteine to form a tetrahedral intermediate stabilized by a canonical oxyanion hole (217). This intermediate then collapses, releasing the amino acid and forming an acyl-enzyme intermediate. In the final step, this intermediate is attacked by water to release the primary bile acid (217, 218). Bile salt hydrolase activity alters host metabolism by poorly defined mechanisms (21), and the amino acid taurine released by these enzymes can be further metabolized by Bilophila wadsworthia, eventually being transformed into the pro-inflammatory end product hydrogen sulfide (68, 219).

Following bile salt hydrolysis, the released bile acid steroid core is subject to a diverse set of further gut microbial transformations, including reductive dehydroxylation. These microbial modifications of produce secondary bile acids, which have a number of biological effects, including direct interactions with the host immune system and contribution to growth inhibition of the intestinal pathogen C. difficile (220, 221). The bile acid pool comprises a vast diversity of different molecules, and the host and microbial targets of these molecules, as well as the gut
microbial organisms and enzymes involved in their metabolism, are subjects of ongoing research (211).

Figure 1.23 Gut microbial metabolism of bile salts and bile acids.
Bile salts are bile acids conjugated to amino acids. These compounds are released into the gastrointestinal lumen to facilitate absorption of fats and insoluble vitamins. Microbial bile salt hydrolases cleave amide bonds to release the amino acid from the bile acid steroid core. The released amino acids can be further metabolized by gut microbes. Microbial transformation of the bile acid steroid core itself produces secondary bile acids, which have diverse effects on the microbial community and the host. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.

1.7.3 Hormones

Microbes also metabolize other steroid-like compounds, including estrogen, a hormone that plays diverse physiological roles in the human body and is a main driver of the development of certain cancers (222, 223). As part of normal estrogen excretion, the host releases glucuronidated estrogens into the gut lumen (75). However, gut microbial β-glucuronidases (GUS) can remove the glucuronide group, releasing the biologically active free estrogen for re-absorption by the host into circulation (Figure 1.24). A critical role for microbes in estrogen cycling has been demonstrated in humans, where antibiotic administration significantly increases excretion of
conjugated estrogen (224). A recent screen by Redinbo and co-workers identified the molecular basis of this metabolism, finding that that 17 out of 35 tested microbial GUS enzymes can de-conjugate estrogen glucuronide in vitro (225). Although the activities vary substantially between different enzymes, kinetic and structural characterization suggested that some enzymes may have evolved specifically for estrogen deconjugation. Microbial GUS inhibitors consistently decreased this metabolism in complex fecal samples, indicating that microbial activation of estrogen is a widespread and GUS-dependent activity. However, the observation that inhibition of microbial GUS activity had no impact on estrogen-dependent cancer development in mice suggests that much remains to be understood about microbiota-estrogen interactions in vivo (225). The microbiota has the potential to interact with many other host hormones (226), including what types, raising the possibility that additional hormones could be subject to microbial transformations.

Figure 1.24 Gut microbial metabolism of hormones.
Gut microbes are critical for the bioactivation of estrogen hormones. For example, gut microbes hydrolyze estradiol-3-glucuronide into free estradiol, the active hormone. Gluc stands for glucuronide. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.7.4 Neurotransmitters

Finally, gut microbes interact with host neurotransmitters. Recent work has implicated the gut microbiota in the development of neurological disease and in the regulation of host neurotransmitter levels (102, 179, 227, 228). Although the precise mechanisms dictating microbiota-nervous system interactions in health and disease remain to be elucidated, one potential mechanism is the direct microbial production and metabolism of neurotransmitters (Figure 1.24). Many different bacteria can produce neurotransmitters in vitro, including γ-aminobutyric acid (GABA), histamine, and tryptamine (178). The biological effects of microbial neurotransmitter synthesis remain poorly understood *in vivo*, but recent work on tryptamine suggests a role in gut motility and the enteric nervous system (229). In addition to biosynthesis, gut microbes directly metabolize neurotransmitters, and this can influence their growth. GABA was recently identified as a growth factor for the bacterium KLE173. This previously uncultured member of the Ruminococcaceae family had long been listed as a priority for isolation on the National Institutes of Health “most wanted list”. Strikingly, this organism did not use any of over 100 other tested nitrogen and carbon sources and lacked genes involved in metabolism of common sugars, suggesting it has evolved to use a neurotransmitter as an essential nutrient (230) (Figure 1.24). Similarly, the neurotransmitter serotonin is a growth factor for the human gut bacterium *Turicibacter sanguinis* and promotes its intestinal colonization (179). This bacterium directly imports serotonin into the cell and although serotonin induces major metabolic and transcriptional changes, the precise mechanisms underlying the growth-stimulating effect are still unclear (179).

The gut microbiota is also involved in metabolism of catecholamines. The GI tract is a major site of host biosynthesis of the dopamine and norepinephrine, key catecholamines that are involved in gut motility and inflammation (231-234). Catecholamines increase the virulence of
intestinal pathogens, helping them acquire iron during inflammation (231, 232), and are also subject to direct metabolism by commensal bacteria. Gut microbes can hydrolyze glucuronide-conjugated catecholamines, releasing the active neurotransmitters into the lumen, and can also directly modify the catecholamine scaffold of dopamine and norepinephrine (Figure 1.24) (80, 186, 235, 236). For example, the gut microbial dehydroxylation of dopamine into m-tyramine is a key step in gut microbial metabolism of the Parkinson’s disease drug L-dopa (237), and the finding that healthy individuals secrete m-tyramine in urine suggests that dopamine dehydroxylation takes place under homeostatic conditions as well (238). The organisms and enzymes that mediate this metabolism, as well as its biological effects on the microbial community and the host, are unknown. A recent metagenomic study identified a significant correlation between predicted genetic pathways for microbial dopamine oxidation and lower rates of depression human patients (239), highlighting that microbiota-catecholamine interactions may play a role in mental health. Importantly, however, this pathway was only predicted computationally based on metabolic pathway modeling, and no gut bacteria has been demonstrated to directly perform this predicted reaction in culture.
Figure 1.25 Gut microbial metabolism of neurotransmitters.
A) Gut microbial decarboxylation of the amino acid produces the neurotransmitter GABA, which is an essential nutrient for the bacterium KLE1738. B) Gut microbial β-glucuronidases hydrolytically release the bioactive neurotransmitter L-norepinephrine in the gut. This can be further metabolized via catechol dehydroxylation into m-octopamine, a metabolite with unknown biological activity. C) Gut microbial β-glucuronidases hydrolytically release the bioactive neurotransmitter dopamine in the gut. This can be further metabolized via catechol dehydroxylation into m-tyramine, a metabolite with unknown biological activity. Dopamine can potentially be oxidized by microbes into DOPAC. The presence of a predicted but not experimentally verified metabolic pathway for microbial dopamine oxidation pathways, as assessed by metagenomics, correlates with lower rates of depression in humans. Known effects of microbial metabolism or properties of metabolites are shown in parentheses. The examples shown are selected microbial pathways rather than an exhaustive and comprehensive list.
1.8 Approaches for gut microbial pathway and enzyme discovery

Attempts to decipher the biological consequences of gut microbial metabolism have been hindered by our poor understanding of these transformations at the molecular level. Although many associations exist between xenobiotic and endogenous compounds and the microbiota, we have limited information about the distribution of these activities in patient populations. Metabolic functions rarely correlate directly with microbial phylogeny, and considerable strain-level variation exists even within the same species (15), which limits the information gained from assessing the composition of the gut microbiota alone. Large-scale metagenomic analyses that claim to detect differences in putative xenobiotic metabolism pathways (240, 241) have not provided experimental evidence supporting predicted changes in activities. To elucidate how gut microbial metabolism influences human health, it is critical to connect functions of interest with genes and enzymes. Below we highlight recent case studies illustrating prominent approaches that enable discovery and manipulation of gut microbial enzymes and metabolic pathways.

1.8.1 Discovering new microbial transformations

A majority of gut microbiota-dependent transformations have been identified on a case-by-case basis through studies of individual molecules, rather than through systematic large-scale screening efforts. Thus, the molecules known to be metabolized may represent those that have historically been deemed interesting. Given the vast metabolic potential of the human gut microbiota, known examples of microbial metabolism likely represent only a small fraction of the scope of potential metabolic transformations carried out by gut microbes. Below we highlight common approaches for discovering novel microbial transformations.
1.8.1.1 High-throughput screening

Recent work illustrates the power of high-throughput screening in discovering new microbiota transformations and addressing the gap between currently known metabolism and the vast metabolic potential of the gut microbiota (Figure 1.26). Recently, Goodman and co-workers performed a systematic screen incubating 76 diverse human bacteria with a pool of 271 orally administered drugs (Figure 1.26). Using Liquid Chromatography-Mass Spectrometry (LC-MS) to detect metabolism, they identified that two-thirds of drugs can be depleted by at least one microbial isolate, revealing many novel, previously unknown metabolic transformations, including degradation of contraceptives and antipsychotics (150). Overall, they found that microbial metabolism generally involved a core set of transformations (reduction, hydrolysis, functional group transfer, as discussed above), suggesting a majority of transformations may be performed by a subset of core protein families.

To identify the enzymes responsible for select transformations, they expressed libraries of gDNA from metabolizing bacteria in E. coli and searched for gain of activity in this heterologous host. Goodman and co-workers initially focused on Bacteroides thetaiotaomicron as a model for this approach, using genetics and in vitro biochemistry to follow up on the gain-of-function assays and confirm 17 gene products metabolizing 18 different drugs, including the calcium channel blocker diltiazem. Metabolism of this of diltiazem was further characterized in gnotobiotic mice colonized with wild-type and mutant B. thetaiotaomicron strains differing only in the presence of the metabolizing enzyme. These experiments revealed that the newly discovered bacterial enzyme was critical for the intestinal deacetylation of diltiazem and its metabolites. Expanding on their gain-of-function approach to the drug-metabolizing bacteria Bacteroides dorei and Collinsella aerofaciens, the authors identified an additional 13 gene products from that transformed 16
different drugs. Notably, the drug-metabolizing gene products identified in *B. thetaiotaomicron, B. dorei* and *C. aerofaciens* were enriched among a collection of diverse gut bacterial species that perform the corresponding metabolism, indicating a broader relevance of these genes in gut microbial drug metabolism. Additionally, presence of drug-metabolizing genes significantly correlated with metabolism of the three tested drugs diltiazem, norethindrone acetate, and famciclovir, in complex fecal microbiota cultures ex vivo. These results suggested the identified genes may be important in a more ecologically relevant context. As a whole, this study provides a framework for discovering new transformations, identifying potential genes and enzymes involved, and linking microbial genes to variability in drug metabolism among individual strains and gut microbiota communities (150).

While the screen by Goodman and co-workers supports the idea that the scope of microbial drug metabolism is far greater than approximately 50 transformations that had been previously reported (149), the work is limited by the small number of strains evaluated relative to the diversity of gut bacteria. In addition, some metabolism may only be present in the context of a complex, native gut microbial community where ecologically relevant microbe-microbe interactions are conserved. Overcoming these limitations, Donia and co-workers recently performed a screen with a complex human fecal sample instead of the library microbial isolates used by Goodman and co-workers (Figure 1.2) (242). Screening 575 different drugs, they identified metabolism of 13% of drugs, representing 28 distinct structural classes and many novel transformations. Because only one human fecal sample was used, it is possible that this number represents only a small fraction of the possible transformations that might be observed a greater diversity of microbiota samples. Overall, the screen uncovered that the gut microbiota metabolized a significantly higher proportion
of naturally derived molecules compared to synthetically derived drugs, suggesting that the gut microbiota is more likely to metabolize molecules and scaffolds that it more frequently encounters.

Donia and co-workers focused on the deglycosylation of the chemotherapeutic capecitabine as a proof-of-principle for further characterization. Genetics implicated a uridine phosphorylase in this metabolism in the model metabolizing organism *E. coli*. To explore the relevance of microbial capecitabine metabolism to drug efficacy in a host setting, the authors conducted a pharmacokinetic experiment in mice. First, they treated conventional mice with an antibiotic cocktail to deplete the endogenous microbiota. Following the microbiota depletion, they colonized a group of the “pseudo germ-free” mice with the capecitabine-metabolizing human fecal sample used in the initial drug screen. Another group of non-colonized mice served as a control. Subsequent pharmacokinetic profiling in serum revealed that the newly discovered gemicitabine glycosylation indeed occurred in vivo and was microbiota-dependent. However, it is unclear whether the identified *E. coli* uridine phosphorylase is responsible for the observed microbiota-dependent effects. Nonetheless, this work represents an intriguing approach for rapid discovery of potential microbial drug metabolism directly from human fecal samples.

Culture-based approaches have been traditionally used to uncover microbial drug metabolism, and the studies by Goodman and Donia and co-workers showcase how high-throughput screen can accelerate the discovery process to identify a greater range of potential drug-microbiota interactions. However, there are some important caveats. For example, screening directly in fecal samples presents possible challenges with organism identification, as a specific transformation that has been identified in a complex fecal sample may not be clearly linked to common laboratory strains such as *E. coli* or other available isolates. In such cases, the metabolizing fecal sample may be plated, and individual microbial colonies screened for
metabolism (172). Given that some transformations carry growth benefits for the metabolizing organism, enrichment using selective growth media could also be used (243). Furthermore, in some cases, apparent drug depletion in cultures could result from drug sequestration rather than direct metabolism (244). Therefore, to identify new metabolic pathways in culture-based settings, it is critical to track both disappearance of starting material and the appearance of metabolites, ideally through the use of labeled substrates. Finally, the activity discovered in cultures with strains or communities may not reflect activity in vivo due to many factors, including exposure to the drug in the gut, the location of metabolizing microbes, and absence of important host-microbe and microbe-microbe interactions that only take place inside the environment of the host. Nonetheless, high-throughput approaches hold promise for discovering the scope of potential drug-microbiota interactions. Although systematic screening efforts have thus far largely focused on drug metabolism, similar efforts could have the potential to identify new transformations of dietary compounds and endogenous molecules.
Figure 1.26 Use of high-throughput screens to discover new microbial transformations.
In this approach, compound libraries are incubated with isolated gut microbial strains or complex gut microbiota (stool) samples. Liquid Chromatography-Mass Spectrometry (LC-MS) is then used to identify compound depletion or production of novel metabolites, revealing potentially novel drug-microbiota interactions. Strains, genes, and enzymes responsible for metabolism can then be identified.
1.8.1.2 Computationally guided discovery

Culture-based screens of individual gut microbial strains and complex microbial communities rely on our ability to culture bacteria in the laboratory. However, many strains in the gut microbiota are unculturable using current methods, raising the possibility that culture-based screens could fail to uncover the full diversity of transformations. Computationally guided enzyme discovery overcomes the limitation of culture-based methods and provide opportunities to prioritize discoveries based on metagenomic abundance or other variables relevant to the ecology of the gut microbiota (Figure 1.27). This is evidenced by a recent study of GREs, an enzyme superfamily that is abundant in the human gut microbiota but harbors many uncharacterized members (25, 65). Combining protein Sequence Similarity Network analysis with quantitative metagenomics, Balskus and co-workers found that some specific uncharacterized GREs were particularly abundant in human gut metagenomes relative to other habitats, making them promising targets for functional characterization and suggesting potentially relevant ecological roles (71) (Figure 1.27). Biochemical assays uncovered an enzyme that dehydrates trans-4-hydroxy-L-proline, an amino acid that is abundant in the diet, is a major component of host collagen, and can serve as an electron acceptor for Clostridia during Stickland fermentation (245). This approach, named chemically guided functional profiling (CGFP) also identified a propanediol dehydratase involved in metabolism of host-derived sugars (71). Similar computational analyses of protein superfamilies have been successfully employed to identify and characterize gut microbial β-glucuronidases (246).

Other computational tools integrate data across diverse datasets to generate hypothesis about novel metabolic activities. For example, the computation tool Microbe-FDT groups vast numbers of drugs, dietary components, and host molecules into clusters based on chemical
similarities (Figure 1.2) (247). Clusters of chemically similar compounds are then linked to information about associated protein families and known toxicities, such that if one molecule in a cluster has a known association with a microbial enzyme and a toxicity, it is assumed that other molecules in that cluster might be metabolized by similar microbial enzymes and be associated with similar toxicities. This tool can predict the known substrate scope the digoxin-metabolizing enzyme cardiac glycoside reductase 2 (Cgr2) and also uncovered a previously unreported demethylation of the ovarian cancer drug Atretamine by complex fecal samples (247). Although microbe-FDT will be valuable tool for hypothesis generation, an important caveat is its reliance on known enzymatic activities and correct functional annotations for accurate predictions. Given that a high percentage of gut microbial proteins cannot be assigned any function, lack homology to known enzymes, and are poorly or incorrectly annotated (10), the utility of computational methods will be significantly aided by fundamental efforts to functionally characterize gut microbial genes and enzymes.
Figure 1.27 Use of computationally guided approaches to discover new microbial transformations.
A) Chemically Guided Functional Profiling (CGFP) first uses Sequence Similarity Network (SSN) analysis to group members of a protein superfamily into putatively isofunctional clusters.
Some clusters do not have biochemically characterized members, indicating members of this cluster have unknown activity. SHORTBRED generates unique marker sequences for each cluster, which can then be used for quantitative metagenomics. This allows assessment of the relative abundance of distributions of potential ecological relevance can be selected for further characterization. Protein expression and biochemistry enables identification of the enzymatic activity, potentially revealing novel metabolic pathways. B) Microbe-FDT groups compounds based on structural similarity, creating clusters with structurally similar molecules that may derive from different sources. Information about toxicity and associated enzyme families are then added to specific compounds on the chemical network. Compounds with known properties (such as compound 1) may be connected to compounds with unknown toxicity and enzyme families on the network (compound 2), suggesting potentially shared properties. Incubation of compound 2 with isolated strains or complex fecal samples could reveal new modes of metabolism and potentially provide insight into microbiota-dependent toxicity.

1.8.2 Linking known activities to genes and enzymes

Once microbial transformations have been linked to specific microbial strains, the next challenge is to identify the specific genes and enzymes involved in metabolism. Below we use case studies to highlight approaches and implications of linking known transformations to genes and enzymes in human gut bacteria.

1.8.2.1 Genetic screens

Genetics is perhaps the most classically used approach for identifying gene functions in bacteria, and recent developments of genetic tools in non-model organisms such as human gut Bacteroides and Clostridia has enabled significant progress in the study of microbial transformations (98, 248). For instance, Goodman and co-workers recently used a genetic screen to identify microbial genes involved in the hydrolysis of the drug Brivudine (BRV) into bromovinyluracil (BVU) (Figure 1.28) (14). A number of abundant gut bacteria metabolize this drug in anaerobic culture, including Bacteroides thetaiotaomicron. Lacking a clear gene candidate in B. thetaiotaomicron, Goodman and co-workers used transposon mutagenesis to disrupt 75% of predicted non-essential genes, and then screened this mutant library for BRV metabolism. This
genetic screen implicated a single locus in metabolism, and subsequent in-frame gene disruption and assays with purified protein confirmed the involvement of a predicted purine nucleoside phosphorylase in this reaction (Figure 1.27). Because the host also hydrolyzes BRV to BVU, identification of the microbial enzyme responsible enabled investigations of the extent to which microbes contribute to BRV metabolism in a simplified gnotobiotic model system. This study highlights the utility of a classical genetic screens for assigning microbial gene functions. However, this approach is limited in scope given that a majority of organisms remain genetically intractable. Advances in CRISPR and genome engineering tools may promise to expand the genetic toolkit to a wider range of gut bacteria, but this remains an important obstacle in understanding gut microbial metabolic activities (249).

![Figure 1.28](image)

**Figure 1.28 Use of genetic screens to link transformations to genes and enzymes.** Genetic approaches were used to characterize genes and enzymes involved in gut microbial Brivudine metabolism. A) The prominent gut bacterium *B. thetaiotaomicron* hydrolyzes the anti-viral drug Brivudine into BVU, a metabolite implicated in lethal drug-drug interactions in humans. B) To identify the genes responsible, random transposon mutagenesis was performed in *B. thetaiotaomicron*. The resulting mutant library was incubated with Brivudine, and subsequent analysis of metabolites in bacterial cultures unveiled a mutant with a loss-of-function phenotype. This mutant carried a disruption in a predicted purine nucleoside hydrolase. Subsequent targeted in-frame knockout and enzyme purification verified the activity.
1.8.2.2 Chemically guided genome mining

Enzyme identification can also be guided by a clear hypothesis of underlying chemistry involved. A prominent example involves microbial choline metabolism. Gut microbes anaerobically convert choline, an essential nutrient found in meat, eggs, and milk, into TMA, which is further transformed into trimethylamine N-oxide (TMAO) by FMOs in the liver (Figure 1.28). This co-metabolic pathway is implicated in several human diseases, including cardiovascular disease \( (104, 105) \), and although microbial choline metabolism had been reported many decades ago, the gut microbial genes responsible were unknown. A chemically guided, rational genome-mining effort ultimately identified the genes and enzymes that mediate this process (Figure 1.28) \( (250) \). Choline fermentation begins with a C–N bond cleavage reaction that resembles the first step in ethanolamine utilization, a transformation catalyzed by a B\(_{12}\)-dependent radical enzyme (ethanolamine ammonia-lyase). Hypothesizing that these pathways might share certain reactions, the Balskus group searched the genome of the choline-metabolizing, animal-associated strain *Desulfovibrio desulfuricans* for homologs of known ethanolamine-degrading enzymes from the pathogen *Salmonella enterica*. This analysis revealed the choline utilization (*cut*) gene cluster and CutC, a GRE that converts choline into TMA and acetaldehyde \( (67, 251, 252) \).

Biochemical and structural studies have revealed CutC active-site residues that interact with the trimethylammonium group of choline and mediate C–N bond cleavage \( (67) \). These conserved amino acids are specific to this particular GRE and have enabled accurate identification of the *cut* pathway in numerous, phylogenetically diverse human gut bacteria \( (243) \). The discovery of CutC has aided interventions to modulate choline metabolism in vivo, including the design of
gut communities with decreased TMA production (252) and small-molecule inhibitors targeting this pathway (253-255).

**Figure 1.29 Use of chemically guided genome mining to link transformations to genes and enzymes.**

Genome mining was used to identify the microbial pathway involved in the microbial transformation of choline to trimethylamine (TMA), a precursor of the disease-associated metabolite trimethylamine N-oxide (TMAO). A) Pathway for microbial and host co-metabolism of choline. B) Microbial choline metabolism resembled ethanolamine metabolism, suggesting potentially shared biochemical logic. Genome mining for acetaldehyde-processing enzymes from the ethanolamine pathway in choline-metabolizing bacteria identified the choline utilization (*cut*) gene cluster. Further biochemical and structural characterization confirmed the involvement of choline TMA-lyase (*CutC*) in the transformation of choline into TMA, enabling opportunities to develop targeted inhibitors and to identify choline-metabolizing gut microbes.

### 1.8.2.3 Transcriptomics and comparative genomics

Genome mining relies on a clear hypothesis for the enzymes responsible for a reaction of interest. However, in some cases genome mining is not viable, either because the chemistry is novel and has no enzymatic precedent, or because a specific transformation could be performed by many candidate examples. The discovery of genes and enzymes that metabolize the cardiac drug digoxin highlights an alternative approach for functional discovery: the use of transcriptomics and comparative genomics. Digoxin has a very narrow therapeutic window, requiring careful monitoring to avoid toxicity. More than 10% of United States patients taking digoxin excrete high
levels of dihydrodigoxin, an inactive metabolite derived from reduction of an α,β-unsaturated lactone (Figure 1.30) (256). Early studies implicated the gut microbiota in drug inactivation, as co-administering digoxin with antibiotics decreased or abolished dihydrodigoxin production (257). Moreover, fecal samples from dihydrodigoxin-excreting individuals were found to completely metabolize digoxin. Subsequent isolation of digoxin-metabolizing microbes revealed that a single organism, *Eubacterium lentum* (renamed *Eggerthella lenta*), was responsible (172). However, *E. lenta* was also found in patients who did not excrete dihydrodigoxin, illustrating that the presence of this species alone was not predictive of activity.

Although *E. lenta*’s role in digoxin metabolism had been appreciated for decades, challenges in growing the organism and a lack of genetic tools hampered efforts to understand this transformation. More recently, Turnbaugh and co-workers used RNA sequencing (RNA-seq) and comparative genomics to identify a digoxin-inducible gene cluster present only in digoxin-metabolizing *E. lenta* strains (Figure 1.30) (18). The cardiac glycoside reductase (*cgr*) operon from this organism encodes two proteins (*Cgr1* and *Cgr2*) that resemble bacterial reductases involved in anaerobic respiration. Because *E. lenta* is genetically intractable, Turnbaugh and co-workers turned to comparative genomics between closely related *E. lenta* strains. Strikingly, the presence of *cgr* genes correlates with digoxin reduction by these strains, providing genetic evidence for its involvement in the reaction. Furthermore, experiments in GF mice mono associated with *E. lenta* strains and incubations with human fecal samples indicate that these genes could be useful biomarkers for digoxin inactivation (18). Follow-up studies with purified enzymes revealed that *Cgr2* is a flavin and [4Fe-4S] cluster-dependent enzyme that directly reduces the lactone ring of digoxin (15). Surprisingly, *Cgr2* possesses a strikingly narrow substrate scope, metabolizing only digoxin and other closely related cardiac cardenolides. This protein represents a novel enzyme
with low homology to any biochemically characterized reductases, suggesting it may have evolved for cardiac glycoside metabolism (15). The studies of digoxin metabolism delineate an approach that is generalizable to additional inducible microbial metabolic activities whose enzymatic basis is unknown.

Figure 1.30 Use of transcriptomics and comparative genomics to link transformations to genes and enzymes.
Researchers used transcriptomics and comparative genomics to identify genes involved in the metabolism of the cardiac drug digoxin by the genetically intractable gut microbe *E. lenta*. A) *E. lenta* reduces the unsaturated lactone of digoxin, producing an inactive drug metabolite. B) RNA-sequencing identified a unique set of genes upregulated in response to digoxin in *E. lenta*. Comparative genomics in metabolizing and non-metabolizing *E. lenta* strains uncovered Cgr1 and Cgr2 as likely candidates involved in the metabolism. These genes predicted metabolism by complex human fecal samples. Cgr2 was later biochemically validated as the digoxin-reducing enzyme.

1.8.2.4 Native enzyme purification from metabolizing organisms

Similar to the discovery of digoxin-metabolizing genes, the identification of enzymes involved in soy isoflavone metabolism was also complicated by the lack of a clear genetic hypothesis and genetic intractability of the host organism. Soy isoflavones are abundant polyphenols that are poorly absorbed in the small intestine and are subject to microbial metabolism. Gut bacteria metabolize daidzein, a major isoflavone component of soy, to either *O-*
desmethylandolensin (O-dma) or equol using two independent pathways (258-261). For both equol and O-dma, the first step appears to be the same: reduction of daidzein into dihydrodaidzein via a conjugate reduction. Whereas O-dma results from a subsequent E1cB-type ring opening reaction, equol is produced through keto reduction and aliphatic dehydroxylation (113) (Figure 1.31). Levels of both of these metabolites differ among individuals, with 30-60% of humans harboring equol-producing microbes and 80-90% possessing O-dma-producing organisms (262). Whereas the biological effects of O-dma are poorly understood, equol has high phytoestrogenic activity (263) and is associated with a decreased risk of hormone-dependent cancers in humans (264). The potentially chemoprotective effects of equol made it an attractive target for further study.

Various strains isolated from human and rat feces had been implicated in the multi-step conversion of daidzein to equol, but our understanding of this metabolism had been limited because enzymes responsible remained unknown until work by Shimada and co-workers (Figure 1.31). In a 2010 study, the authors demonstrated conversion of daidzein to dihydrodaidzein by soluble lysate fractions from Lactococcus sp. 20-92 (265). To identify the enzyme responsible for this reductive transformation, they fractionated the soluble lysates using a variety of chromatography methods, eventually a 70-kDa protein whose intensity correlated with daidzein reduction activity (Figure 1.31). Subsequent proteomics uncovered the sequence of this protein, helping Shimada and identify the encoding gene. The daidzein reductase, which encoded for a 4Fe-4S cluster containing enoate reductase-like protein, was then cloned and expressed in E. coli, further confirming its involvement in the conjugate reduction of daidzein (265). This was the first ever demonstration of an enzyme capable of soy isoflavone metabolism in vitro.

In a follow-up study, Shimada and co-workers identified the enzymes performing the remaining two reductive steps in the daidzein to equol pathway (Figure 1.31) (266). Daidzein to
equol conversion was inducible in *Lactococcus* sp. 20-92, and qPCR assays demonstrated upregulation of the previously reported daidzein reductase gene in response to daidzein. This led to the hypothesis that the daidzein reductase was part of a larger operon harboring all necessary genes for daidzein to equol conversion. By sequencing the genomic region surrounding the daidzein reductase identified in the earlier study, the authors identified two additional reductases predicted to be sufficient to generate equol. Cloning and heterologous expression in *E. coli* biochemically confirmed the role of these reductases in the remaining two steps in the daidzein metabolic pathway (266). Follow-up work identified the racemase missing from the pathway and characterized similar enzymes from other equol producers (267, 268), but it still remains to be seen whether these genes serve as biomarkers for equol production and disease development in humans. Nonetheless, these studies testify to the power of native protein purification for enzyme discovery. Although native purification was more commonly employed before the introduction of DNA cloning and sequencing, this method not only offers a promising avenue for enzyme identification in the absence of a genetic hypothesis but can also overcome challenges associated with heterologous expression (including but not limited to genetic incompatibility between native and heterologous host, need for specific cofactors, and reliance of system-specific chaperones for protein maturation) (15).
Figure 1.31 Use of native purification to link transformations to genes and enzymes. Gut microbial metabolism of daizein to equol illustrates the use of native purification for enzyme identification. A) Human gut microbes convert the abundant soy isoflavone daidzein to the potent phytoestrogen equol. B) *Lactococcus* sp. 20-92 converted daidzein to equol in cultures. The soluble cell lysate converted daidzein to dihydrodaidzein, the first step in the pathway. To identify enzyme responsible, soluble lysates were fractionated, and the resulting fractions were screened for daidzein reduction activity. This effort identified a 70-kDa protein correlating with activity. Proteomics revealed its sequence, enabling genomic analysis. Similar reductases of unknown function co-localized with the daidzein reductase, and subsequent biochemical assays confirmed their involvement in the remaining reductive steps of the equol pathway.
1.8.2.5 Functional metagenomics

Many approaches for gene and enzyme discovery rely on our ability to culture microorganisms, but the use of functional metagenomics circumvents the need for culturing altogether. In functional metagenomics, DNA from an environmental sample (such as a stool sample) is cloned into plasmids and then expressed in a heterologous host. Clones harboring these gene fragments are then screened for gain of function phenotypes, and subsequent DNA sequencing of the clones can uncover the gene and source organisms responsible. This approach has been successfully employed to link diverse phenotypes, including carbohydrate utilization, cofactor transport, antibiotic resistance, and immune modulation, to microbial genes and enzymes (269-274). This approach overcomes challenges associated with the unculturability of many gut organisms but still suffers from the obstacles associated successful expression of enzymes in heterologous hosts.

1.9 The development and use of small molecules to manipulate microbial metabolism

Identification of enzymes and organisms involved in specific transformations not only uncovers potential biomarkers, but also provides opportunities to manipulate metabolic pathways to elucidate their biological impacts on individual organisms, the microbial community, and the host. Moreover, manipulating gut microbial metabolism represents a promising therapeutic avenue because many metabolic gut microbial pathways are associated with human health or disease. Genetics represents one approach for manipulation, allowing researchers to carefully manipulate specific pathways to generate non-metabolizing mutant strains. Colonizing germ-free animals with a non-metabolizing strain and comparing its effects to those of a wild-type strain has uncovered many key new insights into the effects of microbial metabolism on host biology (18, 21, 98). However, genetic tools are still limited in their scope, with many gut bacteria remaining genetically
intractable (275). Moreover, it remains challenging to genetically manipulate specific pathways in
the context of an intact, complex microbial community, and so most studies are done with isolated
strains or model communities in the absence of native ecological interactions. In contrast to
genetics, small-molecule inhibitors offer a promising avenue for pathway manipulating genetically
intractable bacteria and targeting specific metabolic pathways across diverse metabolizing
bacteria. Small-molecule inhibitors of gut microbial enzymes could uncover the biological roles
of particular transformations and serve as excellent starting points for developing microbiota-
targeted drugs. Below we use case studies to review approaches for the development and use of
small-molecule inhibitors in the context of human gut microbiota.

1.9.1 Discovering leads through high-throughput screening of compound libraries

One of the earliest examples of small-molecule inhibition of microbial metabolism
involved the anti-cancer prodrug Irinotecan (CPT-11). As described in section 1.6.3, the microbial
reactivation of the active drug SN-38 in the gastrointestinal tract causes intestinal damage and
severe diarrhea, side effects that limit the use of this otherwise effective drug (Figure 1.32) (17).
This made inhibition of gut bacterial β-glucuronidases an intriguing approach for preventing drug
reactivation and limiting toxicity.

Because bacterial β-glucuronidases are broadly distributed in commensal bacteria and are
present in humans, any small molecules inhibitors need to be selective for bacterial β-
glucuronidases and nontoxic to both host cells and other gut microbes. Redinbo and co-workers
used an in vitro high-throughput screen of >10,000 compounds to successfully identify potent and
selective inhibitors of purified E. coli β-glucuronidase (Figure 1.32) (17). These inhibitors were
also effective against multiple, distantly related gut bacteria but did not target the human β-
glucuronidase. Structural studies revealed that these compounds interact with an active-site loop
distinct to bacterial β-glucuronidases, which explains their selectivity (17). Administration of one of these inhibitors to mice prevented reactivation of SN-38 in the gut and concomitant toxicity. Further work examining the crystal structures and activities of additional gut bacterial β-glucuronidases identified a conserved Asn-Lys motif that interacts with the carboxylic acid of the glucuronic acid sugar. This motif is absent from glycosidases that accept different substrates, which will help to identify these enzymes in sequencing data sets and elucidate the effects of inhibitors on different types of gut organisms (36). There has been some follow up work that has identified some interesting additional inhibitors, including slow-binding inhibitors that have to be modified by the enzyme itself, which might be neat to mention. Because bacterial β-glucuronidases can deconjugate glucuronides derived from many dietary compounds and drugs, inhibitors of these enzymes may be useful in other therapeutic contexts (41).
Figure 1.32 Use of high-throughput screening to identify microbiota-targeted inhibitors.

Microbial re-activation of the chemotherapy drug irinotecan represents one of the earliest examples of development and use of microbiota-targeted small molecule inhibitors. A) Host metabolism converts irinotecan into SN-38, the active chemotherapeutic. This metabolite is glucuronidated and excreted into the gut lumen, where microbial β-glucuronidases can re-activate the drug, contributing to dose-limiting diarrhea in patients. B) High-throughput screening of purified *E. coli* β-glucuronidase with a >10,000-membered compound library revealed potent inhibitors preventing β-glucuronidase activity towards a fluorescent probe. The lead inhibitor was potent, non-toxic, and selective for microbial β-glucuronidases. The inhibitor ameliorated microbiota-dependent drug toxicity in mice.

1.9.2 Rational inhibitor design

Rational design is another promising avenue for inhibitor development (Figure 1.33). In this approach, biochemical and structural characterization of enzymes guides design of small molecules that interfere with the catalytic cycle, outcompete the native substrate, or inhibit the enzyme by other means. For example, the structure of CutC, a key gut microbial enzyme in disease-associated choline metabolism, enabled identification of betaine aldehyde, a substrate analog that binds to a conserved active site cysteine and forms a covalent thiohemiacetal adduct in the unactivated enzyme (Figure 1.33) (254). Betaine aldehyde is non-toxic to metabolizing organisms and inhibits choline metabolism in complex fecal samples, indicating it might serve as a tool to study this metabolism in vivo (254). Separate efforts have identified additional substrate
analogs that inhibit microbial choline metabolism, and these inhibitors can alleviate some of the deleterious effects of microbial choline metabolism in vivo (253, 255). More recently, Devlin and coworkers used structural information to develop a specific inhibitor of gut microbial bile-salt hydrolases (276). A bile acid derivative containing an α-fluoromethyl ketone electrophilic warhead was found to covalently modify an active site cysteine residue, thus intercepting the catalytic cycle (Figure 1.33). This inhibitor is non-toxic to gut microbes and mammalian cells, inhibits microbial metabolism by diverse strains and in complex fecal samples, and decreases the quantity of free bile acids in conventional mice (276). Inhibitors like these offer promising tools for elucidating the biological consequences of microbial metabolism. Importantly, however, rational design necessitates detailed knowledge of the target, making it difficult to apply to poorly characterized enzymes or enzyme families.
Figure 1.33 Use of rational design to identify gut microbiota-targeted small-molecule inhibitors.

A) The inhibition of choline TMA-lyase (CutC) represents another example of rational inhibitor design in the context of the gut microbiota. D) Structural analysis of the CutC active site revealed residues and substrate features important for substrate recognition and catalysis.
This enabled the identification of betaine aldehyde (BA), a small molecule that covalently modifies an active-site cysteine and displays dose-dependent inhibition of TMA production in complex human fecal samples. B) The microbial hydrolysis of bile salts into primary bile acids represented an intriguing target for small-molecule inhibition. C) Structural analysis and biochemical experiments enabled rational design of a bile-acid like molecule carrying an $\alpha$-fluoromethyl ketone warhead that alkylates an active-site cysteine residue. The inhibitor candidate blocks metabolism by isolated strains and complex gut microbial communities.

1.9.3 Evaluating host-targeted drugs towards gut microbial targets

In addition to developing novel small molecules to manipulate microbial metabolism, many drugs already prescribed in the clinic may serve as promising inhibitor candidates. Hundreds of FDA-approved drugs directly impact gut microbial growth in vitro, and pharmaceutical drug intake correlates with microbial community structure in patient cohorts (153, 277-279). Recent work supports the idea that these drugs may target microbial metabolism (Figure 1.34). For example, the FDA-approved drug raloxifene can inhibit microbial $\beta$-glucuronidases involved in the reactivation of regorafenib, a cancer drug associated with adverse gastrointestinal side effects (280). Moreover, acarbose, a human alpha-amylase inhibitor taken by diabetes patients, inhibits growth of gut microbial Bacteroides species on potato starch and the oligosaccharide pullulan (281). Strikingly, the drug only inhibits growth on polysaccharides that require the bacterial Sus system for extracellular oligosaccharide hydrolysis and uptake, making the effect of acarbose dependent on the microbial nutrient source. Consistent with this, acarbose changes microbial community composition in a diet-dependent fashion in mice (282). Finally, the anti-depressant and selective serotonin reuptake inhibitor (SSRI) fluoxetine (also known as Prozac) inhibits serotonin uptake by the gut microbe T. sanguinis, potentially by targeting a serotonin transporter (179). Fluoxetine also alters T. sanguinis gene expression, inhibits its serotonin-induced growth in culture, and significantly decreases its colonization in antibiotic-treated mice (179). Although the
precise mechanisms underlying these effects is unclear, these results may explain the independent observation that SSRI use is associated with reduced *Turicibacteraceae* in humans (279), highlighting a broader role for psychopharmacological agents in modulating microbial community composition.

Taken together, the surprising findings described above suggest a strategy of repurposing host-targeted drugs to manipulate gut microbial metabolism and indicates that some drugs may have previously unappreciated effects on gut microbes. As more information about drug-microbiota interactions is emerging (153), it will become important to understand how drugs impact microbial transformations relevant to human health and disease. In cases where host and gut microbiota are capable of the same transformation (14, 237), it will be particularly important to evaluate how drugs targeting a host pathway impact the corresponding microbial pathway.
Figure 1.34 FDA approved drugs can inhibit gut microbial enzymes.
A) Some FDA-approved drugs with known human targets may have unappreciated targets in the human gut microbiota, representing new opportunities for small-molecule inhibition of microbial metabolism. B) The osteoporosis drug raloxifene potently inhibits the microbial re-activation of the chemotherapeutic regorafenib. C) The anti-diabetic drug acarbose inhibits microbial growth on oligosaccharides that require the starch uptake system (Sus). D) the antidepressant and selective serotonin reuptake inhibitor fluoxetine (also known as Prozac) blocks serotonin uptake by the gut microbe *T. sanguinis*, resulting in altered gene expression and decreased growth and intestinal colonization.
1.10 Chapter preview

This thesis describes work performed to understand microbial metabolism of dietary compounds, drugs, and host neurotransmitters. Chapter 2 describes our efforts to characterize and inhibit microbial enzymes involved in the metabolism of L-dopa, the main treatment for Parkinson’s disease. We used genome mining, genetics, and biochemistry to identify and characterize a promiscuous tyrosine decarboxylase from the human gut microbe Enterococcus faecalis that directly decarboxylates L-dopa into dopamine. In addition to exploring the distribution of L-dopa decarboxylation among gut bacterial and environmental strains, we used culture-based assays to characterize the relationship between the tyrosine decarboxylase gene and L-dopa metabolism by complex human fecal samples from Parkinson’s patients and neurologically healthy individuals. These assays revealed the existence of alternative pathways for microbial L-dopa metabolism, and we find that the gut bacterium Clostridium sporogenes uses characterized enzymes involved in aromatic amino acid metabolism to degrade L-dopa. Finally, we evaluated the activity of known host decarboxylase inhibitors in microbial L-dopa decarboxylation and rationally identified a small molecule (α-fluoromethyltyrosine) that selectively inhibits microbial L-dopa decarboxylation by pure strains and complex microbial communities.

Chapter 3 describes our work identifying and characterizing microbial strains, genes, and enzymes involved in the dehydroxylation of dopamine, a human neurotransmitter and the intermediate in the two-step microbial metabolism of L-dopa into m-tyramine. We used enrichment culturing to isolate a strain of Eggerthella lenta that dehydroxylates dopamine into m-tyramine. We then used RNA-sequencing, comparative genomics, chemical genetics, and in vitro biochemistry to identify a molybdenum-dependent enzyme (named dopamine dehydroxylase, or Dadh) that mediates this reaction. Bioinformatics analysis revealed that a single nucleotide
polymorphism in the gene encoding this enzyme correlated with metabolism by isolated strains and complex microbial communities.

Chapter 4 expands on our discovery of Dadh. Dopamine dehydroxylation is just one of many examples of catechol dehydroxylation, a prominent transformation in the microbial metabolism of endogenous compounds, dietary small molecules, and FDA-approved drugs. Chapter 4 catalogs our efforts to understand the molecular basis of catechol dehydroxylation in the human gut microbiota. First, we characterized the substrate scope and regulation of the dopamine dehydroxylase and find a remarkable specificity for the catecholamine scaffold. We also demonstrated that dopamine can serve as an alternative electron acceptor in *E. lenta*, promoting growth of metabolizing strains in pure cultures and in defined and complex microbial communities. To identify strains capable of metabolizing catechols other than dopamine, we performed a colorimetric screen of gut bacteria incubated with catecholic drugs, dietary compounds, and siderophores. We found that *Gordonibacter pamelaeae* 3C dehydroxylates DOPAC and closely related *E. lenta* dehydroxylated hydrocaffeic acid, and (+)-catechin. RNA-sequencing and comparative genomics revealed candidate molybdenum-dependent enzymes involved in these reactions. The native purification of the hydrocaffeic acid dehydroxylase (Hcdh) strongly supported a model in which each catechol substrate is processed by its own dedicated enzyme. Phylogenetic and sequence similarity network analyses established catechol dehydroxylases as a new class of molybdenum-dependent enzymes and demonstrate that unknown members of this class are widespread among microbes in the environment. Finally, we found that catechol dehydroxylation is present among gut microbiotas of mammals representing diverse diets and phylogenetic origins, reinforcing that this activity can take place in habitats beyond the human gut.
Finally, Chapter 5 describes our progress towards further biochemically characterizing catechol dehydroxylases. Based on the Dadh and Hcdh substrate scope and phylogenetic analysis presented in Chapter 4, we proposed a mechanism for catechol dehydroxylation. Preliminary deuterium incorporation experiments investigating hydrocaffeic acid dehydroxylation cell lysates supported this mechanistic proposal, which represents a new strategy for aromatic ring dehydroxylation. To obtain higher yield and purity of catechol dehydroxylases for further biochemical and mechanistic characterization, we turned to heterologous expression. We describe our efforts cloning and expressing the dopamine dehydroxylase (Dadh), hydrocaffeic acid dehydroxylase (Hcdh), catechin dehydroxylase (Cadh), DOPAC dehydroxylase (Dodh), and catechol lignan dehydroxylase (Cldh) in heterologous hosts. These efforts involve extensive testing of different plasmids across different heterologous and different media conditions. No constructs yielded soluble, active protein, suggesting further work is necessary to overcome the challenges associated with the complex maturation and assembly of molybdenum-dependent enzymes.
1.11 References


49. F. Rafii, J. D. Hall, C. E. Cerniglia, Mutagenicity of azo dyes used in foods, drugs and cosmetics before and after reduction by Clostridium species from the human intestinal tract. *Food Chem Toxicol* 35, 897-901 (1997).


190. R. P. Anderson, T. J. Butt, V. S. Chadwick, Hepatobiliary excretion of bacterial formyl-

Metabolism of azetirelin, a new thyrotropin-releasing hormone (TRH) analogue, by intestinal 


Young, S. E. Jones, J. P. Davignon, A clinical trial of amygdalin (Laetrile) in the treatment of 


195. G. J. Strugala, A. G. Rauws, R. Elbers, Intestinal first pass metabolism of amygdalin in the 

Transforming berberine into its intestine-absorbable form by the gut microbiota. *Sci Rep* 5, 12155 

197. R. N. Alolga, Y. Fan, Z. Chen, L. W. Liu, Y. J. Zhao, J. Li, Y. Chen, M. D. Lai, P. Li, L. W. 
Qi, Significant pharmacokinetic differences of berberine are attributable to variations in gut 

198. C. E. Wagner, K. M. Wheeler, K. Ribbeck, Mucins and Their Role in Shaping the Functions 

199. H. Schneider, T. Pelaseyed, F. Svensson, M. E. V. Johansson, Study of mucin turnover in the 

J. I. Gordon, Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science* 307, 


Chapter 2. Discovery, characterization, and inhibition of gut microbial enzymes involved in the metabolism of L-dopa, the main treatment for Parkinson’s disease

This chapter was adapted in part from:
2.1 Introduction

As outlined in Chapter 1, microbial interactions with drugs are implicated in human health and disease, including neurological disease. A growing body of work links the trillions of microbes that inhabit the human gastrointestinal tract (the human gut microbiota) to neurological conditions, including Parkinson’s disease, a debilitating neurodegenerative disorder that affects more than 1% of the global population over 60 (1-3). Gut microbes from Parkinson’s patients exacerbate motor deficits when transplanted into germ-free mouse models of disease, and gut microbial proteins directly contribute to pathophysiology (2, 4). In addition, multiple studies have revealed differences in gut microbiota composition in Parkinson’s disease patients compared with healthy controls that may correlate with disease severity (5-12). These observations suggest that the gut microbiota plays a role in neurodegeneration and support the long-standing hypothesis that Parkinson’s – generally thought of as a disease of the brain – may start in the gut (2, 4, 13-19). However, the influence of the human gut microbiota on the treatment of Parkinson’s and other neurodegenerative diseases remains poorly understood due to the lack of information about the molecular mechanisms involved.

The primary treatment for Parkinson’s disease is Levodopa (L-dopa) (20), an amino acid prescribed to manage motor symptoms that result from dopaminergic neuron loss in the substantia nigra. After crossing the blood-brain barrier, L-dopa is decarboxylated by aromatic amino acid decarboxylase (AADC) to give dopamine, the active therapeutic agent (Figure 2.1). However, only 1 to 5% of L-dopa reaches the brain, owing to extensive pre-systemic metabolism that mainly takes place in the gut (21-25). Several human enzymes metabolize L-dopa, including catechol-O-methyltransferase (COMT), but AADC is the major enzyme involved in the extensive degradation of L-dopa in the gastrointestinal tract (Figure 2.1) (26-30). Peripheral decarboxylation is
problematic not only because it decreases drug bioavailability, but also because it produces bioactive dopamine. Peripherally produced dopamine cannot cross the blood-brain barrier, and the accumulation of dopamine can lead to gastrointestinal side effects, orthostatic hypotension through activation of vascular dopamine receptors, and cardiac arrhythmias (31, 32). To decrease peripheral metabolism, L-dopa is generally co-administered with AADC inhibitors such as carbidopa and benserazide as well as COMT inhibitors (Figure 2.1). Nonetheless, even in the presence of these inhibitors, peripheral metabolism remains a problem, with as much as 56% of L-dopa still being metabolized peripherally in the presence of carbidopa, the major decarboxylase inhibitor used in the United States (33). Furthermore, patients display highly variable responses to L-dopa, including high interindividual variability in side effects and loss of efficacy over time (34). Such variation between patients and their response to medication is well-recognized but remains poorly understood.
Figure 2.1 Mechanism of action and host metabolism of L-dopa.
Taken orally, L-dopa crosses the blood-brain barrier and reaches the brain where the human AADC enzyme transforms it into dopamine, the active therapeutic agent. However, AADC also decarboxylates a substantial portion of L-dopa peripherally in the gastrointestinal tract, limiting bioavailability and producing dopamine that does not cross the blood-brain barrier and instead causes side effects. Because AADC is the major host enzyme involved in peripheral metabolism, AADC inhibitors such as carbidopa and benserazide are often co-prescribed with L-dopa. These compounds limit peripheral decarboxylation and help more of the drug reach the brain.

Multiple lines of evidence suggest that gut microbial interactions with L-dopa influence treatment outcomes. Parkinson’s disease drugs are associated with gut community composition independent of disease status, suggesting direct interactions between these drugs and the microbiota (8). In addition, administering broad-spectrum antibiotics improves L-dopa therapy, suggesting that gut bacteria interfere with drug efficacy (35, 36). One mode of interference could be microbial metabolism of L-dopa, which could reduce L-dopa bioavailability and lead to side effects (37-40). Interestingly, knowledge of gut microbial L-dopa metabolism dates back over six decades. In 1957, Jones and co-workers observed that rabbits and rats fed L-dopa excreted the
metabolite \textit{m}-hydroxyphenylacetic acid (\textit{m}-hpaa) in urine, providing the first evidence of a pathway that would later be implicated in microbial metabolism (37). Follow-up work identified variable levels of this metabolite in urine of six different patients taking L-dopa, with as much as 1% of the total L-dopa dose accounted for by \textit{m}-hpaa (41). Importantly, \textit{m}-hpaa levels significantly decreased significantly following antibiotic (neomycin) treatment (41). Separately, it was observed that rat caecal contents directly could degrade L-dopa (42). L-dopa was never labeled in these initial studies, making it difficult to establish a direct route from L-dopa to \textit{m}-hpaa. However, the effects of antibiotics on the metabolism in humans, the evidence of direct L-dopa metabolism by rat cecal contents, and the knowledge that only gut microbes (and not the host) could perform the catechol dehydroxylation reaction required to produce \textit{m}-hpaa (43), strongly implicated gut microbes in L-dopa metabolism.

Studies in patients and rodents elucidated a major microbial pathway involved in L-dopa metabolism. Similar to \textit{m}-hpaa, the metabolite \textit{m}-tyramine was excreted in patients receiving L-dopa and depleted by neomycin treatment, leading to the hypothesis that \textit{m}-tyramine might be a precursor for \textit{m}-hpaa (39). The available data at the time presented two possible routes for gut microbial production of L-dopa from \textit{m}-hpaa. In one scenario, L-dopa would be directly metabolized to DOPAC, which would be dehydroxylated to yield \textit{m}-hpaa. In the other scenario, L-dopa would be decarboxylated to dopamine, which would be dehydroxylated to \textit{m}-tyramine, a metabolite that the host would oxidize to \textit{m}-hpaa using known pathways for amine oxidation (Figure 2.2). Each case involved a proposed, exclusively microbial catechol dehydroxylation reaction. Evidence for both pathways was observed when rat cecal contents were anaerobically incubated with L-dopa ex vivo (38), making it unclear which pathway was the major contributor in the host.
Comparing urinary metabolites between conventional and germ-free rats fed L-dopa and potential intermediate metabolites, Goldman and co-workers revealed that a majority of excreted \( m \)-hpaa derives from host oxidation of \( m \)-tyramine \((38, 44)\). In their initial animal experiments, rats were fed L-dopa through the diet, making it difficult to establish the percentage of consumed substrate that was converted to a particular metabolite. This limitation notwithstanding, their studies allowed for critical comparisons that shed light on the key steps involved in microbial L-dopa metabolism. First, they found that conventional rats fed 100 mg L-dopa or 100 mg dopamine as part of their diet produced comparable levels of urinary \( m \)-hpaa (2.3 mg and 4.0 mg over a 72-hour period, respectively). No \( m \)-hpaa was produced from these molecules in germ-free animals, reinforcing the previous idea that \( m \)-hpaa requires microbial metabolism for its production. In contrast, when conventional rats received 100 mg DOPAC, the precursor that was predicted to be dehydroxylated to yield \( m \)-hpaa, only 0.3 mg of \( m \)-hpaa was recovered in urine over 72 hours, indicating that the dehydroxylation of DOPAC into \( m \)-hpaa does not quantitatively account for the \( m \)-hpaa levels observed in response to L-dopa or dopamine.

Follow-up experiments supported the hypothesis that the major dehydroxylation reaction in the L-dopa pathway was the transformation of dopamine into \( m \)-tyramine rather than the transformation of DOPAC into \( m \)-hpaa \((38, 45)\). Conventional rats fed dietary dopamine accumulated urinary DOPAC rapidly, with peak levels reached within one day of feeding and no changes on the subsequent day of feeding. In contrast, higher levels of \( m \)-hpaa accumulated on the second day of feeding, suggesting dehydroxylation had a unique pharmacokinetic profile decoupled from DOPAC production. Additionally, food-deprivation, which was used to starve the gut microbiota and manipulate its metabolic activity, significantly decreased urinary \( m \)-hppa in conventional rats fed dopamine, from 6 mg on day 1 of collection to 2.4 mg on day 4 \((45)\). Urinary
DOPAC levels did not decrease, indicating that the microbiota likely does not contribute majorly to this metabolite (45). Finally, both germ-free and conventional animals converted dopamine to DOPAC and also oxidized m-tyramine to m-hpaa (38). Based on these metabolite data, Goldman and co-workers proposed that the major microbial metabolic pathway involved decarboxylation of L-dopa into dopamine, a step that mirrors host metabolism, followed by a uniquely microbial dehydroxylation of the catechol ring of dopamine to produce m-tyramine (Figure 2.2) (38, 45, 46).

Subsequent comprehensive profiling of L-dopa and its host and microbial urinary metabolites in conventional rats suggested that more than 5% of orally ingested L-dopa is metabolized through dopamine into m-tyramine (46). However, the precise contribution of gut microbial L-dopa decarboxylation remains unclear, as both host and microbes decarboxylate L-dopa and all microbially generated dopamine may not be dehydroxylated into m-tyramine. Additionally, it is worth noting that early studies of rat cecal incubations and conventional and germ-free rats revealed the existence of a third possible pathway for microbial L-dopa metabolism, resembling anaerobic tyrosine fermentation and involving an additional catechol dehydroxylation to form m-hydroxyphenylacetic acid (m-hpaa) (Figure 2.2). This pathway makes only minor contributions to drug metabolism in vivo, with an average of 0.12 mg appearing in urine of conventional rats fed L-dopa over a three-day period, compared to as much as 3.8 mg of m-hpaa in the same animals (38, 47).
Figure 2.2 Gut microbial metabolism of L-dopa.
The major gut microbial pathway involves decarboxylation of L-dopa into dopamine, followed by dehydroxylation into \( m \)-tyramine, a metabolite that is oxidized by the host into \( m \)-hydroxyphenylacetic acid (\( m \)-hpaa). Alternatively, L-dopa can be metabolized by an unknown pathway into DOPAC, followed by gut microbial dehydroxylation. Additionally, an unknown microbial pathway transforms L-dopa into hydrocaffeic acid, followed by catechol dehydroxylation to produce \( m \)-hydroxyphenylpropionic acid (\( m \)-hppa). However, these two latter pathways are thought to make only minor contributions to drug metabolism in vivo.

No microbial isolate from any habitat has ever been demonstrated to dehydroxylate dopamine, but previous work has shed light on strains that decarboxylate L-dopa. For example, cell suspensions and freeze-dried cells from a strain of Streptococcus faecalis were reported to decarboxylate L-dopa into dopamine and the related aromatic amino acid tyrosine into \( p \)-tyramine (48-51). The same enzyme was thought to perform the decarboxylation of tyrosine and L-dopa in this organism, but there was initially controversy surrounding this claim (52). S. faecalis inhabits the human gut but also diverse environments on the planet, and the uncertain origin of the S. faecalis strain used in these early studies made the relevance of this activity to the human gut
microbiota unclear (53). *Streptococcus faecalis* was renamed *Enterococcus faecalis* many decades after these initial observations based on genetic methods, and the enzyme responsible for tyrosine decarboxylation in *E. faecalis* was subsequently genetically identified (54, 55). Biochemical assays with purified *E. faecalis* and related tyrosine decarboxylases revealed the requirement of the cofactor pyridoxal-L-phosphate (PLP) and characterized the activity towards tyrosine and phenylalanine (56-59). The structural similarity between L-dopa and tyrosine suggest that a tyrosine decarboxylase might be responsible for L-dopa decarboxylation in *E. faecalis*; however, this possibility was never evaluated with purified enzyme, and studies of PLP-dependent aromatic amino acid decarboxylases from food-associated and soil-associated microbes have revealed varying substrate preferences, challenging this idea. For example, whereas a decarboxylase from *Lactobacillus brevis* IOEB 9809 is specific for tyrosine and a decarboxylase from *Pseudomonas putida* only turns over L-dopa, a tyrosine decarboxylase from *L. brevis* CGMCC 1.2028 is promiscuous for L-dopa (58-61).

In summary, at the time we began our work, the microbial genes and enzymes responsible for L-dopa decarboxylation in the human gut microbiota remained unknown. The clinical relevance of this pathway was also unclear given the potential effects of co-administered inhibitors of host peripheral L-dopa metabolism on these gut microbial activities. Below we describe our efforts to understand the molecular basis of L-dopa decarboxylation in the human gut microbiota.
2.2 Results

2.2.1 Discovery of strains that decarboxylate L-dopa

We sought to elucidate the genetic and biochemical bases for gut microbial L-dopa metabolism and understand how co-administered AADC inhibitors affect this pathway. Using a genome mining approach, we first identified strains encoding candidate L-dopa decarboxylating enzymes. Aromatic amino acid decarboxylation is typically performed by enzymes employing PLP an organic cofactor that provides an electron sink (Figure 2.3) (62). As described above, the PLP-dependent tyrosine decarboxylase (TyrDC) from the food-associated strain *Lactobacillus brevis* CGMCC 1.2028 was shown to have promiscuous activity toward L-dopa in vitro. To our knowledge, this was the only purified and biochemically characterize microbial aromatic amino acid decarboxylase to have this promiscuity, making a promising target sequence for searching in microbial genomes (59). To locate TyrDC homologs in human gut bacteria, we (together with Dr. Jordan Bisanz, Turnbaugh lab, UCSF) performed a BLASTP search against the complete set of Human Microbiome Project (HMP) reference genomes available via NCBI. The majority of TyrDC homologs (above a cutoff of 29% amino acid ID, 78% query coverage) were found in the neighboring genus *Enterococcus*, with some hits within lactobacilli and Proteobacteria (Figures 2.4 and 2.5). To evaluate these hits for potential activity, we selected 10 representative gut strains containing TyrDC homologs (29-100% amino acid ID) and examined their ability to decarboxylate L-dopa in anaerobic culture. Both *Enterococcus faecalis* and *Enterococcus faecium* decarboxylated this drug, suggesting that this activity is present among *Enterococci* (Figure 2.6). To further examine this activity among *Enterococci*, we expanded our screen to include additional *Enterococci* isolates and related strains from diverse environments (53). TyrDC homologs (>60% amino acid identity) were present only in specific clades but correlated with activity across
phylogeny (Figures 2.7 and 2.8). Altogether, these data implicate TyrDC in L-dopa
decarboxylation by Enterococci, including human gut bacteria.

**Figure 2.3 General decarboxylation mechanism of PLP-dependent decarboxylases.**
A conserved lysine residue covalently binds to the PLP cofactor and is later replaced by the amine of the amino acid. Decarboxylation and isomerization and hydrolysis releases the primary amine and generates the cofactor.
Figure 2.4 Phylogenetic distribution of tyrosine decarboxylases (TyrDC) in the human microbiota.

Human Microbiome Project reference genomes were queried by BLASTP for homologs of the *L. brevis* TyrDC, and the results are visualized on a cladogram of phylogeny (based on 16S rRNA alignment). TyrDC homologs found sporadically within *Lactobacillus* spp. (*Lb*) are widely distributed amongst *Enterococcus* (*Ec*; average amino acid identity 67.8% over 97.6% query length. The analysis was performed by and the figure was made by Dr. Jordan Bisanz in the lab of Peter Turnbaugh at UCSF.
Figure 2.5 Phylogenetic tree based on the alignment of the TyrDC protein sequences from Enterococcal HMP reference genomes.

The maximum likelihood phylogenetic tree of TyrDC was prepared using FastTree and shows that the TyrDC sequences correlate with species (abbreviated Spp). Bootstrap values are shown at each node of the tree. The analysis was performed by Dr. Jordan Bisanz in the lab of Peter Turnbaugh at UCSF.

Figure 2.6 Testing representative gut microbial strains encoding TyrDC reveals that E. faecalis strains reproducibly convert L-dopa to dopamine.

Strains were cultured for 48 hours anaerobically. Conversion from L-dopa to dopamine was measured by LC-MS. Bar graphs represent the mean ± the standard error of the mean (SEM) of three biological replicates.
Figure 2.7 Presence of tyrosine decarboxylase homologs across Enterococci and related strains.
The tree represents the expanded phylogeny of the phylogeny presented in Gilmore and co-workers (2016) (53). An orange dot indicates presence of a TyrDC homolog (>60% amino acid identity). The tree was constructed by Dr. Francois Lebreton in the laboratory of Michael Gilmore at Harvard Medical School.
Figure 2.8 L-dopa decarboxylation activity among Enterococci and related strains. Strains were cultured for 48 hours anaerobically. Bar graphs represent the mean ± the SEM of three biological replicates. Metabolism was measured using LC-MS. A red dot indicates presence of a TyrDC homolog (>60% amino acid identity).
2.2.2 The human gut bacterium Enterococcus faecalis decarboxylates L-dopa using a promiscuous PLP-dependent tyrosine decarboxylase

We chose *E. faecalis*, a prevalent, genetically tractable gut organism, as a model for characterizing L-dopa decarboxylation in further detail (53). Importantly, *E. faecalis* was the only gut microbial strain displaying quantitative L-dopa decarboxylation among all strains tested (Figure 2.6). These strains share the highly conserved four-gene *tyrDC* operon (Figure 2.9), and Dr. Jordan Bisanz (Turnbaugh lab, UCSF) found *tyrDC* in 98.4% of the *E. faecalis* assemblies deposited in NCBI with a median amino acid identity of 99.8 (range 97.0-100). This high degree of sequence conservation and prevalence is consistent with tyrosine decarboxylation being a common phenotypic trait of *E. faecalis* and suggest L-dopa metabolism may be broadly present across strains (63).

We used genetics and in vitro biochemistry experiments to confirm that TyrDC is necessary and sufficient for L-dopa decarboxylation by *E. faecalis*. Four different *E. faecalis* MMH594 mutants carrying a 2 kb Tet-cassette disrupting *tyrDC* could not decarboxylate L-dopa (Figure 2.10) and displayed no growth defects compared to wild-type (Figure 2.11). These transposon mutants were generated and provided by Dr. Elizabeth Selleck in the laboratory of Michael Gilmore at Harvard Medical School. We cloned *E. faecalis* MMH594 into the pET28a vector and expressed the protein in *E. coli* BL21 (DE3). In vitro characterization of the heterologously expressed and purified TyrDC revealed a 5-fold higher catalytic efficiency towards L-tyrosine compared to L-dopa, suggesting drug metabolism arises from promiscuous enzyme activity (Figure 2.12 and Table 2.1). Although lyophilized *E. faecalis* cells decarboxylate L-dopa (50) and the *tyrDC* operon’s role in tyrosine decarboxylation in *E. faecalis* is well-characterized (54), the connection between *tyrDC* and L-dopa decarboxylation was unknown. In addition, TyrDC from
*E. faecalis* was previously shown to decarboxylate tyrosine and phenylalanine *(56, 57, 64, 65)*, but its ability to accept L-dopa had not been demonstrated. The apparent selectivity of TyrDC contrasts sharply with that of the human L-dopa decarboxylase AADC, which displays very low activity towards L-tyrosine *(66)*.

**Figure 2.9 Genomic organization of the tyrDC operon in *E. faecalis* strains used in this study.**

A tBLASTn search was performed in the NCBI nucleotide collection and whole genome shotgun sequences using *E. faecalis* TX0104 tyrDC as the query (UniProt ID: C0X244), and the genomic context surrounding the tyrDC gene was analyzed. This search revealed that all *E. faecalis* strains used in our study share the same general four-gene organization of the tyrDC operon. Homologous genes are highlighted in the same color. The text inside each gene represents the NCBI accession number for the protein coding sequence.
Figure 2.10 *E. faecalis* MMH594 *tyrDC* mutants do not decarboxylate L-dopa.

A) PCR confirmation of mutants. gDNA from *E. faecalis* OG1RF (lane 1), *E. faecalis* TX0104 (lane 2), *E. faecalis* MMH594 (lane 3), or *E. faecalis* MMH594 *tyrDC* mutants (lanes 4-7) were amplified using primers surrounding the location of the predicted 2 kb tetracycline resistance cassette insertion. While all wild-type *E. faecalis* strains displayed a 500 kb amplicon representing wild-type *tyrDC*, the mutant displayed a 2.5 kb amplicon, representing a 2 kb insertion. B) L-dopa metabolism by mutants. Strains were cultured for 48 hours anaerobically. Bar graphs represent the mean ± the standard error of the mean (SEM) of three biological replicates. Mutnat number three was used for further studies.
Figure 2.11 Anaerobic growth of wild-type (WT) *E. faecalis* MMH594 and a tyrDC mutant. Strains were grown anaerobically in BHI medium at 37 °C either without L-dopa (A) or with 500 µM L-dopa (B). WT is shown in black, while the mutant is shown in red. There was no obvious growth defect of the *tyrDC* mutant under these conditions. The data shown are the mean of three replicate growth experiments ± standard error of the mean (SEM).

Figure 2.12 The *E. faecalis* tyrosine decarboxylase decarboxylates L-dopa and tyrosine. A) SDS-PAGE of purified recombinant TyrDC from *E. faecalis* MMH594. Ladder is the Precision Plus Protein™ All Blue Standards (first lane from the left). B) Kinetic analysis of *E. faecalis* TyrDC reveals a preference for tyrosine. Error bars represent the mean ± the SEM of three biological replicates.
Table 2.1 Michaelis–Menten parameters determined for TyrDC.
Data represent the best-fit values and their associated standard error (n=3 replicates).

<table>
<thead>
<tr>
<th></th>
<th>Tyrosine</th>
<th>L-dopa</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{\text{max}}$ (μM/s)</td>
<td>6.363 ± 0.102</td>
<td>5.553 ± 0.165</td>
</tr>
<tr>
<td>$K_M$ (μM)</td>
<td>314.9 ± 15.5</td>
<td>1475 ± 79.28</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (1/s)</td>
<td>63.63 ± 1.02</td>
<td>55.53 ± 2.98</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (1/μM*s)</td>
<td>2.02 ± 0.14 *10^5</td>
<td>3.77 ± 0.23 *10^4</td>
</tr>
</tbody>
</table>

2.2.3 Influence of pH and tyrosine concentration on gut microbial L-dopa decarboxylation

Our biochemical studies revealed that tyrosine is the preferred substrate for TyrDC. This substrate is available in the small intestine where a majority of L-dopa decarboxylation takes place, raising questions about the clinical relevance of the promiscuous tyrosine decarboxylase activity. Thus, we next investigated the extent to which tyrosine could interfere with L-dopa decarboxylation by *E. faecalis* (67, 68). In competition experiments, purified TyrDC (Figure 2.13) and anaerobic *E. faecalis* cultures decarboxylated 0.5 mM L-dopa and 0.5 mM tyrosine simultaneously (500 μM tyrosine, approximating the resting small intestinal concentration) (68) (Figure 2.14). This observation sharply contrasts with previous investigations of phenylalanine decarboxylation, which is metabolized by *E. faecalis* only when tyrosine is completely consumed (64). Simultaneous decarboxylation of L-dopa and tyrosine also occurred in *E. faecalis* MMH594 cultures containing higher tyrosine concentrations (1.5 mM, approximating the small intestinal post-meal concentration) (Figure 2.15) and in three human fecal suspensions (Figure 2.16). As observed previously for tyrosine by other groups, L-dopa decarboxylation occurred more rapidly at lower pH across all strains tested (Figures 2.14 and 2.15), suggesting this metabolism is likely accelerated at the lower pH of the upper small intestine (69, 70). As the $K_m$ of TyrDC for L-dopa (1.5 mM) is below the estimated maximum in vivo small intestinal L-dopa concentration even at its lowest clinically administered dose (5 mM), these data strongly suggest that peripheral decarboxylation is performed by both host and gut bacterial enzymes.
Figure 2.13 L-dopa and tyrosine competition experiment with TyrDC.

TyrDC (0.15 µM) was incubated with equimolar concentrations of tyrosine and L-dopa (500 µM each) in 0.2 M pH 5.5 sodium acetate buffer at room temperature. Formation of the corresponding decarboxylation products was measured by LC-MS/MS following quenching with methanol (1:10) at specific time points. The data shown is the mean of three replicate experiments ± the SEM. Error bars are not visible if they are smaller than the data points.
Figure 2.14 Anaerobic growth, L-dopa metabolism, and tyrosine metabolism of four *E. faecalis* strains across varying pH.

*E. faecalis* MMH594, V583, TX0104, and OG1RF were grown anaerobically with 1 mM L-dopa in BHI medium (pH 5 or pH 7) at 37 °C. The BHI medium contained approximately 500 µM of tyrosine. Growth (A) and metabolites (B) were tracked over time. Though different strains displayed variability in the rate of metabolism, L-dopa and tyrosine decarboxylation occurred simultaneously regardless of the pH. The rate of decarboxylation increased at lower pH despite less growth under this condition. The data shown are the mean of three replicate growth experiments ± the SEM. Error bars are not visible if they are smaller than the data points.
Figure 2.15 Impact of pH and tyrosine concentration on L-dopa metabolism by *E. faecalis* MMH594 grown in BHI medium.

*E. faecalis* was grown anaerobically with 1 mM L-dopa in BHI medium (pH 5 or pH 7) at 37 °C. The medium was supplemented either with no tyrosine or 1 mM tyrosine, generating final concentrations of approximately 500 µM tyrosine and 1.5 mM tyrosine, respectively. Growth (A) and metabolites (B) were tracked over time. While higher tyrosine concentrations reduced the rate of L-dopa metabolism, L-dopa and tyrosine decarboxylation occurred simultaneously across all conditions. The rate of decarboxylation increased at lower pH despite less growth under this condition. The data shown are the mean of three replicate growth experiments ± the SEM. Error bars are not visible if they are smaller than the data points.
Figure 2.16 Impact of pH and tyrosine concentration on d₃-phenyl-L-dopa metabolism by three human fecal samples grown in BHI medium

Fecal samples were grown with 1 mM d₃-phenyl-L-dopa in BHI medium (pH 5 or pH 7) at 37 °C. The pH 7 medium was supplemented either with no tyrosine or 1 mM tyrosine, creating final concentrations of approximately 500 µM tyrosine and 1.5 mM tyrosine, respectively. Growth (A) and metabolites (B) were tracked over time. While tyrosine and pH had variable effects depending on the stool sample, L-dopa and tyrosine decarboxylation occurred simultaneously across all samples and conditions. The data shown are the mean of three replicate growth experiments ± the SEM. Error bars are not visible if they are smaller than the data points.
2.2.4 L-dopa decarboxylation promotes survival of *E. faecalis* in acidic conditions

We next investigated the physiological effects of L-dopa metabolism in *E. faecalis*. Previous work suggests that tyrosine decarboxylation promotes survival of lactic acid bacteria such as *E. faecalis* in acidic environments (69). The PLP-dependent amino acid decarboxylation mechanism involves the irreversible addition of a proton to the substrate, and the export of this substrate effectively removes this proton from the cell, making the cell interior less acidic and generating a proton motive force (69, 71-73). This de-acidification mechanism helps lactic acid bacteria such as *Enterococci* survive in acidic fermented foods and potentially promotes their survival during passage through the stomach (74, 75). To evaluate the potential involvement of L-dopa in this survival mechanism, we measured the growth of *E. faecalis* in a minimal, acidic medium (pH 5) in the presence and absence of the drug. L-dopa significantly promoted growth of the wild-type strain over a vehicle control (Figure 2.17). This was due to active decarboxylation, as supplementation of dopamine itself did not promote growth over vehicle. Moreover, the growth of the *tyrDC* mutant remained unchanged in response to L-dopa (Figure 2.17). These data are consistent with previous work on bacterial decarboxylation of tyrosine and highlights the intriguing possibility that L-dopa consumption could impact *E. faecalis* physiology in acidic parts of the GI tract (70, 76).
Figure 2.17 L-dopa decarboxylation promotes acid survival of *E. faecalis*. *E. faecalis* wild-type or a tyrDC mutant were grown in a minimal medium at pH 5 and growth was tracked over time. The data shown are the mean of three replicate growth experiments ± the SEM. Error bars are not visible if they are smaller than the data points.

2.2.5 TyrDC abundance correlates with L-dopa metabolism ex vivo

To investigate whether *E. faecalis* transforms L-dopa in the human gut microbiota, we assessed the metabolism of deuterated L-dopa by ex vivo cultures of stool samples from 19 unrelated neurologically healthy individuals. We used LC-MS to measure metabolism, quantifying both dopamine and *m*-tyramine metabolites produced from L-dopa decarboxylation. As a reminder, *m*-tyramine is typically the endpoint for microbial L-dopa metabolism. While 7/19 samples did not show detectable depletion of L-dopa, the remaining samples displayed significant variability in metabolism, ranging from partial (25%) to almost full conversion (98%) of L-dopa to *m*-tyramine (Figure 2.18). We next asked whether the abundance of tyrDC predicted metabolism in these samples. qPCR enumeration of tyrDC (51) and *E. faecalis* discriminated metabolizing and non-metabolizing samples (p<0.0001, one-tailed Mann-Whitney test) (Figure 2.19). We found a strong linear correlation between tyrDC abundance and *E. faecalis* qPCR abundance in these samples (R² = 0.99, p<0.0001) (Figure 2.20) which likely reflects the high conservation of tyrDC in *E. faecalis* genomes. These data also suggest that *E. faecalis* is the dominant microorganism
responsible for L-dopa decarboxylation in these complex human gut microbial communities. Consistent with this, Dr. Jordan Bisanz (Turnbaugh lab, UCSF) found that *E. faecalis* abundance significantly correlated with *tyrDC* abundance in 1870 human gut microbiomes ($R^2>0.812$, $p<2.2e-16$ Pearson’s Correlation) (Figure 2.21).
Figure 2.18 Metabolism of $d_3$-phenyl-L-dopa by gut microbiota samples from 19 unrelated neurologically healthy individuals ex vivo.
Samples were cultured anaerobically with $d_3$-phenyl-L-dopa (1 mM) for 72 hours and were analyzed by LC-MS/MS. Results are mean concentration ± the SEM (n=3 replicates).

Figure 2.19 tyrDC and E. faecalis abundance correlates with L-dopa metabolism status.
A) The abundance of tyrDC in human gut microbiota samples incubated with 1 mM $d_3$-phenyl-L-dopa. Data represent the average tyrDC abundance (as assessed by qPCR) across the 3 replicates for samples in Figure 2.18. Results are mean ± the SEM. **** P<0.0001, one tailed Mann-Whitney test.

B) Abundance of E. faecalis (as assessed by qPCR) in human gut microbiota samples incubated with 1 mM $d_3$-phenyl-L-dopa. Each data point is the average abundance across 3 biological replicates for each sample shown in Figure 2.18. Results are mean ± the SEM. **** P<0.0001, one tailed Mann-Whitney test.
Figure 2.20 Correlation between tyrDC and *E. faecalis* abundance in ex vivo cultures incubated with L-dopa

Data represent qPCR with *E. faecalis* or tyrDC-specific primers. Each data point is the average abundance across 3 biological replicates for each sample shown in Fig. 3B. Linear regression indicated a highly significant correlation ($R^2 = 0.99$, $p<0.0001$), suggesting *E. faecalis* is likely responsible for decarboxylation in these samples.

Figure 2.21 Correlation between tyrDC and *E. faecalis* abundance in human gut metagenomes.

Correlation of gene/bacterial abundances clearly demonstrates a strong linear correlation between *Enterococcus/tyrDC* ($R^2>0.812$, $P<2.2e-16$ Pearson’s Correlation). The red data points represent organism/enzyme correlation for an unrelated pair. The analysis was done by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
While we found a strong correlation between L-dopa decarboxylation and tyrDC abundance in the ex vivo incubations from these neurologically healthy human subjects, we also noticed that in some samples L-dopa was depleted without a corresponding amount of dopamine and m-tyramine formation, indicating an incomplete mass balance. Previous work in rats has revealed the existence of alternative L-dopa pathways producing phenolic acids from L-dopa ex vivo (38), suggesting these other metabolites might account for the missing masses. These metabolites could conceivably be generated by known fermentative pathways for L-tyrosine and L-phenylalanine, two aromatic amino acids that structurally resemble L-dopa (77). Analyzing a subset of the samples (n=12) further by LC-MS, we detected the presence of deuterated DOPAC and its dehydroxylated metabolite m-hydroxyphenylacetic acid (m-hpaa) (Figure 2.21). We did not detect m-hydroxyphenylpropionic acid, a microbiota-derived L-dopa pathway previously observed in rat urine and rat stool cultures ex vivo (38, 42). In some samples, these additional detected metabolites did not add up to the total amount of L-dopa, suggesting other potential mechanisms of drug depletion. Thus, further work is necessary to determine the factors underlying the mass imbalance.

![Figure 2.22 Detection of alternative L-dopa metabolites in ex vivo cultures from 12 unrelated neurologically healthy human subjects.](image_url)

Samples were cultured anaerobically with d3-phenyl-L-dopa (1 mM) for 72 hours and were analyzed by LC-MS/MS. Results are mean concentration ± the SEM (n=3 replicates).
To further explore the clinical relevance of these findings, we assessed the metabolism of L-dopa by fecal suspensions from Parkinson’s disease patients ex vivo (n = 12). Similar to control subjects, these individuals displayed significant variability in metabolism of L-dopa, with metabolism being present in subjects taking L-dopa/carbidopa at the time of stool sample collection and in drug-native patients (Figure 2.23). qPCR assays revealed that tyrDC abundance and *E. faecalis* abundance discriminated between L-dopa decarboxylating and non-decarboxylating samples (p<0.005, one-tailed Mann-Whitney test) (Figure 2.24). Similar to our ex vivo experiments with stool samples from neurologically healthy patients, we also observed depletion of L-dopa without corresponding production of dopamine or *m*-tyramine in three samples (Figure 2.23). We further analyzed these samples by LC-MS and found that L-dopa was converted to hydroxyphenylpropionic acid (hppa) (Figure 2.23), a pathway that we did not detect in neurologically healthy subjects but is thought to make a minor contribution to drug metabolism in vivo (38, 45, 46). Overall, these data establish a link between tyrDC and L-dopa decarboxylation in Parkinson’s patients and suggest a role for gut bacteria in the extensive interindividual variability in L-dopa decarboxylation observed in Parkinson’s patients (23).
Figure 2.23 Metabolism of $d_3$-phenyl-L-dopa by stool samples from 12 unrelated Parkinson’s disease patients.
A) LC-MS/MS was used to quantify metabolism of $d_3$-phenyl-L-dopa by 12 unrelated human gut microbiota samples from Parkinson’s disease patients ex vivo. Samples were cultured anaerobically in MEGA medium containing $d_3$-phenyl-L-dopa (1 mM) for 72 hours. Metabolite levels in culture supernatants were analyzed by LC-MS/MS. Stacked bar plots represent the mean concentration ± the SEM of three biological replicates. Samples 1-6 represent patients not currently taking L-dopa/carbidopa, while samples 7-12 are from patients taking L-dopa/carbidopa at the time of collection. B) LC-MS/MS was used to quantify phenolic acid metabolites of $d_3$-phenyl-L-dopa in patients showing detectable L-dopa depletion without production of dopamine or $m$-tyramine in panel (A).

Figure 2.24 tyrDC and E. faecalis abundance correlates with L-dopa metabolism status in Parkinson’s patients.
A) The abundance of E. faecalis predicts L-dopa decarboxylation in human gut microbiota samples from Parkinson’s disease patients.
Figure 2.24 (Continued).
Data represent qPCR with \(E. \text{faecalis}\)-specific 16S rRNA primers. Each data point is the average abundance across 3 biological replicates for each sample shown in Figure 2.23A. ** P=0.0040, one tailed Mann-Whitney test. B) The abundance of \(\text{tyrDC}\) predicts L-dopa decarboxylation in human gut microbiota samples from Parkinson’s disease patients. Data represent qPCR with \(\text{tyrDC}\)-specific primers. Each data point is the average abundance across 3 biological replicates for each sample shown in Figure 2.23A. Bars represent the mean and standard error. ** P=0.0020, one tailed Mann-Whitney test.

2.2.6 \(\text{Clostridium sporogenes}\) metabolizes L-dopa using enzymes involved in aromatic amino acid fermentation

Having established the molecular basis of L-dopa decarboxylation in \(\text{E. faecalis}\) and a potential physiological role for this transformation, we next turned our attention to other, orthogonal, pathways that could potentially be involved in gut microbial L-dopa metabolism. The metabolites \(m\)-hpaa and \(m\)-hppa that we observed in our ex vivo incubations likely derive from catechol dehydroxylation of 3,4-dihydroxyphenylacetic acid (DOPAC) and hydrocaffeic acid, respectively (38, 45). These intermediates resemble the end-products of previously described pathways involved in gut microbial fermentation of aromatic amino acids, including phenylalanine and tyrosine (78-80). For example, the prevalent gut microbe \(\text{Clostridium sporogenes}\) metabolizes the amino acid tyrosine, which differs from L-dopa only in the absence of a meta-hydroxyl group, into the metabolites \(p\)-hydroxyphenylpropionic acid (\(p\)-hppa) and \(p\)-hydroxyphenylacetic acid (\(p\)-hpaa) (77). In this recently characterized pathway, tyrosine is first converted into a central 4-hydroxyphenyl pyruvate intermediate by an unknown aminotransferase. This intermediate then undergoes either reductive or oxidative metabolism. In the reductive pathway, the hydroxyphenyl pyruvate is first reduced to 4-hydroxyphenyl lactate. The lactate metabolite is subject to an energetically demanding radical transformation to form 4-hydroxyphenyl acrylate, which is further reduced by conjugate addition into \(p\)-hppa (Figure 2.25) (77). In the orthogonal oxidative branch of the pathway, the central 4-hydroxyphenyl pyruvate intermediate is oxidized into \(p\)-hpaa as the
final product (Figure 2.25). Because of the structural similarity between tyrosine and L-dopa, we hypothesized that these reductive and oxidative tyrosine metabolism pathways could also accept L-dopa.

Figure 2.25 Pathway for tyrosine metabolism by C. sporogenes and proposed pathway for alternative L-dopa metabolism.
A) C. sporogenes transforms tyrosine into 4-hydroxyphenyl pyruvate. This central intermediate then undergoes either reductive or oxidative metabolism by known enzymes (77). B) Proposed alternative pathway for e L-dopa metabolism. C. sporogenes could generate hydrocaffeic acid or DOPAC, which could be further dehydroxylated by unknown members of the gut microbial community.

To test whether the C. sporogenes could metabolize L-dopa, we incubated wild-type C. sporogenes and isogenic strains harboring mutations in the genes encoding metabolic enzymes involved in the two L-tyrosine metabolizing pathways (from the Sonnenburg lab at Stanford University) with deuterium-labeled L-dopa in anaerobic cultures. LC-MS analysis of these cultures showed that C. sporogenes reductively metabolizes L-dopa into hydrocaffeic acid, a metabolite that could potentially be dehydroxylated further by the microbiota into m-hppa (Figure 2.26). Importantly, however, we did not detect DOPAC in this experiment, indicating that the oxidative pathway does not accept L-dopa or that this pathway was not operative under the experimental conditions that we used (Figure 2.26). Further work is necessary to elucidate the origin of these metabolites, including assessing whether tyrosine is processed via oxidative metabolism under our
conditions. It is also possible that DOPAC and its dehydroxylated metabolite \( m \)-hpaa derive from the oxidative tyrosine pathway in a gut microbial strain other than \emph{C. sporogenes}, or from direct de-amination of either dopamine or \( m \)-tyramine. Taken together, our observations provide a plausible explanation for the formation of \( m \)-hppa in our ex vivo experiments and represent the first demonstration of gut microbial strains and enzymes metabolizing L-dopa in this way. Although these pathways make only minor contributions to L-dopa metabolism in vivo (38, 45), their distribution and clinical relevance among patients remains poorly understood.
Figure 2.26 *C. sporogenes* metabolizes L-dopa to hydrocaffeic acid.

A) Proposed metabolites and genes involved in *C. sporogenes* metabolism of *d*₃-phenyl-L-dopa. B-F) mass peak area of metabolites detected upon incubation of *C. sporogenes* wild-type and mutant strains with *d*₃-phenyl-L-dopa. Samples were cultured anaerobically with *d*₃-phenyl-L-dopa (1 mM) for 72 hours in TYG medium and were analyzed by LC-MS/MS. Results are mean concentration ± the SEM (n=3 replicates). Genes downstream of metabolism affected enzymes upstream, consistent with downstream reductions energetically driving metabolic flux (77).
2.2.7 Evaluation of human aromatic amino acid decarboxylase inhibitors towards gut microbial L-dopa decarboxylation

L-dopa is often co-prescribed either with carbidopa or benserazide, two decarboxylase inhibitors that target peripheral AADC in order to increase L-dopa bioavailability. Carbidopa, an L-dopa mimic and the main inhibitor used in the United States, intercepts the catalytic cycle by forming a covalent hydrazone linkage with the PLP cofactor (Figure 2.27) (81, 82). Benserazide is a pro-drug that undergoes hydrolysis to release a serine and unmask a hydrazine warhead that also covalently binds to PLP (Figure 2.27) (83, 84). These inhibitors have the potential to directly interact with the gut microbiota, and the shared general PLP-dependent mechanism of AADC and TyrDC suggests they could inhibit microbial L-dopa decarboxylation (85-88). Consistent with this, we found that the free benserazide hydrazine potently inhibited both human AADC and TyrDC (IC$_{50}$ = 0.36 µM and IC$_{50}$ = 0.59 µM, respectively) and also inhibited L-dopa decarboxylation by growing _E. faecalis_ cultures (EC$_{50}$ = 136.5 µM) (Figure 2.28 and Table 2.2). However, benserazide, the drug that is orally ingested by patients, only weakly inhibited _E. faecalis_ decarboxylation (50% inhibition at 10 mM) and did not display a clear dose-dependent effect on decarboxylation by TyrDC (Figure 2.28). The precise location of benserazide hydrolysis as well as the enzymes responsible for this reaction in the human body is uncertain, making it unclear whether the gut microbiota encounters the drug in its pro-drug or active form. Thus, we focused our further characterizations on carbidopa, an inhibitor that does not require host metabolism for activity.

Testing carbidopa’s activity toward purified enzymes and bacterial cells, we found this drug was 200-fold less active toward purified _E. faecalis_ TyrDC (IC$_{50}$ = 57 µM) relative to _H. sapiens_ AADC (IC$_{50}$ = 0.21 µM) and showed only ~50% inhibition of L-dopa decarboxylation by
*E. faecalis* cultures at the solubility limit of 2 mM (Figure 2.28 and Table 2.2). Additionally, carbidopa did not affect growth of *E. faecalis* (Figure 2.29). Given the maximum predicted gastrointestinal concentration of carbidopa (0.4–9 mM), these data suggest this drug may not fully inhibit gut bacterial L-dopa decarboxylation in Parkinson’s patients. Indeed, we found that 2 mM carbidopa did not alter the kinetics of L-dopa degradation (Figure 2.30) or endpoint *m*-tyramine production in stool samples from both Parkinson’s patients and neurologically healthy controls (Figure 2.31) These observations are consistent with previous findings that carbidopa administration does not impact *m*-tyramine production in patients (89). The low efficacy of carbidopa in whole cells and complex communities may result from poor bacterial cell permeability combined with reduced inhibitory activity for TyrDC.

**Figure 2.27 Human AADC inhibitors and their mechanism of action.**
A) Carbidopa is the decarboxylase inhibitor used in the United States. Benserazide is used in other parts of the world. This inhibitor requires hydrolytic release to unmask the hydrazine warhead. B) Both of these inhibitors form a stable hydrazone linkage with the PLP cofactor, intercepting the catalytic cycle.
Figure 2.28 Effect of benzerazide, benzerazide free hydrazine, and carbidopa on L-dopa decarboxylation by AADC, TyrDC and E. faecalis.
A-C) Activity towards purified enzymes AADC and TyrDC. AADC or TyrDC were incubated with inhibitor and reaction rates were measured by LC-MS/MS. % Activity represents the rate relative to a no inhibitor (vehicle) control. Results are mean ± the SEM (n=3 replicates). A) represents carbidopa, B) represents benzerazide free hydrazine, C) represents the benzerazide prodrug (which was only tested towards the E. faecalis TyrDC. D) Activity of inhibitors towards L-dopa decarboxylation by E. faecalis MMH594 cultures. All data shown are the mean of three replicate growth experiments ± the SEM.
Table 2.2 Inhibition parameters determined for carbidopa, benserazide, and benserazide free hydrazine.
IC$_{50}$ and EC$_{50}$ values (in µM) represent the best-fit values and their associated standard error (n=3 replicates).

<table>
<thead>
<tr>
<th></th>
<th>Carbidopa</th>
<th>Benserazide</th>
<th>Benserazide free hydrazine</th>
</tr>
</thead>
<tbody>
<tr>
<td>TyrDC</td>
<td>52 ± 0.06</td>
<td>No observable dose-dependent inhibition</td>
<td>0.59 ± 0.01</td>
</tr>
<tr>
<td>AADC</td>
<td>0.21 ± 0.01</td>
<td>Not tested</td>
<td>0.36 ± 0.02</td>
</tr>
<tr>
<td><em>E. faecalis</em></td>
<td>50% inhibition at solubility limit of 2 mM</td>
<td>50% inhibition at 10 mM</td>
<td>136.5 ± 0.01</td>
</tr>
</tbody>
</table>

Figure 2.29 Carbidopa does not inhibit growth of *E. faecalis*.
*E. faecalis* MMH594 was grown anaerobically in BHI medium at 37 °C with varying concentrations of carbidopa. There was no obvious dose-dependent growth defect of *E. faecalis* in the presence of this inhibitor. The data shown are the mean of three replicate growth experiments ± the SEM.
Figure 2.30 Time course of degradation of \(d_3\)-phenyl-L-dopa in the presence and absence of 2 mM carbidopa in a stool sample from a neurologically healthy patient.
A metabolizing human fecal sample was incubated in MEGA medium containing \(d_3\)-phenyl-L-dopa (1 mM) with or without carbidopa (2 mM) for 72 hours. LC-MS/MS was used to quantify metabolites at 12, 24, 48, and 72 hours. The data shown are the mean of three replicate experiments ± the SEM.

Figure 2.31 Activity of carbidopa in human fecal microbiota samples.
Fecal samples were incubated anaerobically in MEGA medium containing 0 or 2 mM carbidopa and 1 mM \(d_3\)-phenyl-L-dopa for 72 hours at 37 °C. Metabolites in culture supernatants were analyzed by LC-MS/MS. Error bars represent the mean ± the SEM for three biological replicates. Each panel represents a different individual. Sample B) is from a neurologically healthy patient, while A), C) and D) are from Parkinson’s disease patients.
2.2.8 The small-molecule inhibitor (S)-α-fluoromethyltyrosine selectively targets bacterial decarboxylation

The surprising result that carbidopa only weakly inhibited bacterial L-dopa decarboxylation highlighted the possibility of developing microbial decarboxylase inhibitors. Such molecules could be useful tools to elucidate the microbial contributions to overall metabolism in vivo and could also have therapeutic relevance, as preventing microbial L-dopa decarboxylation could potentially improve L-dopa efficacy. In addition to the hydrazine functional groups found in carbidopa and benserazide, many different groups are known to inhibit PLP-dependent decarboxylases when appended to the amino acid core, including α-monofluoromethyl, α-difluoromethyl, α-vinylmethyl, and α-methyl substituents (Figure 2.32) (88, 90-93). A majority of these groups are warheads derail catalysis by forming a covalent linkage with the cofactor. However, studies of AADC suggest that α-methyl substituents derail catalysis by “tricking” the enzyme into catalyzing oxidative cleavage, producing reactive aldehydes that covalently modify active site residues (94). Decarboxylase inhibitors carrying the α-fluoromethyl warhead are already used in the clinic (for example in the anti-parasitic drug eflornithine) (88), making these molecules promising targets for the development of novel microbial L-dopa decarboxylase inhibitors. Additionally, molecules carrying this warhead are ‘suicide’ inhibitors: inhibitory activity requires that the enzyme deacrbxylates the substrate, offering a layer of selectivity and providing opportunities to target decarboxylases based on their substrate preference (Figure 2.32) (66, 88).
Figure 2.32 Examples of warheads used to inhibit PLP-dependent decarboxylases and mechanism of action of \( \alpha \)-fluoromethyl warheads.

A) Common warheads used to inhibit PLP-dependent decarboxylases. The amino acid substrate is shown as a reference. B) Mechanism of action of \( \alpha \)-fluoromethyl warheads. These molecules undergo active decarboxylation by the enzyme to intercept the catalytic cycle, eventually forming a covalent inhibitor-PLP adduct that can be detected by LC-MS.
To design a selective inhibitor of L-dopa decarboxylation, we first surveyed the substrate scope of TyrDC and AADC. This screen revealed that TyrDC requires a $p$-hydroxyl group for robust activity while AADC prefers a $m$-hydroxyl substituent (Figure 2.33). This led us to hypothesize the L-tyrosine analog (S)-$\alpha$-fluoromethyltyrosine (AFMT) might selectively inhibit the microbial enzyme (Figure 2.34). We received this compound under a Material Transfer Agreement from Merck where it had been synthesized by Kollonitsch and co-workers in the 1970s (66, 88). This molecule had been evaluated as an inhibitor of mammalian AADC in vitro and as a probe for tyrosine hydroxylase in rats but had never been evaluated towards a microbial tyrosine decarboxylase (66, 95, 96).

Consistent with our prediction, AFMT strongly inhibited L-dopa decarboxylation by TyrDC ($IC_{50} = 4.7 \mu M$) but not AADC (~20% inhibition at solubility limit of 650 $\mu M$) in vitro (Figure 2.34 and Table 2.3). Consistent with this selectivity, AFMT formed a covalent PLP adduct only in the presence of TyrDC (Figure 2.35). AFMT was also effective at inhibiting L-dopa metabolism by in E. faecalis cultures ($EC_{50} = 1.4 \mu M$) (Figure 2.36), outperforming carbidopa by 1000-fold without affecting growth (Figure 2.36 and tables 2.2 and 2.3). The fluorine group was essential for potent activity in E. faecalis, supporting that the hypothesis that the mechanism of inhibition in whole cells likely results from covalent decarboxylase inhibition (Figure 2.34). Finally, we found that AFMT completely inhibited L-dopa decarboxylation in gut microbiota samples from Parkinson’s disease patients and neurologically healthy control subjects (Figure 2.37) and was non-toxic to eukaryotic cells (Figure 2.38). Taken together, these data establish AFMT is a selective, non-toxic inhibitor that inhibits microbial L-dopa decarboxylation in complex communities.
Figure 2.33 Screen of amino acid substrates for AADC and TyrDC.
AADC (0.20 µM) or TyrDC (0.15 µM) were incubated with varying substrates (500 µM of L-dopa, phenylalanine, p-tyrosine, or m-tyrosine) at room temperature (0.1 M pH 7.4 phosphate buffer for AADC, 0.2 M pH 5.5 sodium acetate buffer for TyrDC) and reaction rates were measured by LC-MS/MS. The AADC reaction was quenched at 180 seconds and while the TyrDC reaction was quenched at 60 seconds by dissolving 1:10 in methanol. Reactions were performed at room temperature. Reaction supernatants were analyzed for production of the primary amines resulting from decarboxylation by LC-MS/MS. % Activity represents the rate relative to the most rapidly consumed substrate (m-tyrosine for AADC, p-tyrosine for TyrDC). The data shown is the mean of three replicate experiments ± the SEM.
Figure 2.34 Effect of AFMT on L-dopa decarboxylation by AADC, TyrDC and *E. faecalis*.  
A) Structure of AFMT. B) Activity of AFMT and α-methyltyrosine (AMT), towards L-dopa decarboxylation by *E. faecalis* MMH594 cultures. The requirement for the fluorine group for potent activity supports the hypothesis that the mechanism of inhibition in whole cells involves covalent modification of TyrDC. All data shown are the mean of three replicate growth experiments ± the SEM.  
C) Inhibitory activity of AFMT and AMT towards purified enzymes AADC and TyrDC. AADC or TyrDC were incubated with inhibitor and reaction rates were measured by LC-MS/MS. % Activity represents the rate relative to a no inhibitor (vehicle) control. Results are mean ± the SEM (n=3 replicates). AFMT was obtained from Merck under a Material Transfer Agreement.
Table 2.3 Inhibition parameters determined for AFMT.

IC$_{50}$ and EC$_{50}$ values (in µM) represent the best-fit values and their associated standard error (n=3 replicates).

<table>
<thead>
<tr>
<th></th>
<th>AFMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TyrDC</td>
<td>4.71 ± 0.01</td>
</tr>
<tr>
<td>AADC</td>
<td>20% inhibition at solubility limit of 700 µM</td>
</tr>
<tr>
<td><em>E. faecalis</em> MMH594 cells</td>
<td>1.41 ± 0.01</td>
</tr>
</tbody>
</table>

Figure 2.35 Covalent inhibition of *E. faecalis* TyrDC by AFMT.

Detection of an AFMT-PLP covalent adduct following incubation of TyrDC or AADC with AFMT for 1 hour. The data shown is the EIC of mass for the predicted covalent adduct. The mechanism of formation of this adduct is described in Figure 2.32.
Figure 2.36 AFMT does not inhibit growth of *E. faecalis*.
*E. faecalis* MMH594 was grown anaerobically in BHI medium at 37 °C with varying concentrations of AFMT. There was no obvious dose-dependent growth defect of *E. faecalis* in the presence of this inhibitor. The data shown are the mean of three replicate growth experiments ± the SEM.

![Graph showing growth inhibition by AFMT](image)

Figure 2.37 Activity of AFMT in human fecal microbiota samples.
Fecal samples were incubated anaerobically in MEGA medium containing 0 or 2 mM carbidopa and 1 mM *d*3-phenyl-L-dopa for 72 hours at 37 °C. Metabolites in culture supernatants were analyzed by LC-MS/MS. Error bars represent the mean ± the SEM for three biological replicates. Each panel represents a different individual. Samples A), B), and E) are from Parkinson’s disease patients. Samples C) and D) are from neurologically healthy patients.
**Figure 2.38 MTS assay evaluating AMFT toxicity toward HeLa cells.**

HeLa cells were seeded in 100 µL growth medium ([DMEM medium supplemented with 10% FBS (2 mL) and 1X Antibiotic-Antimycotic (100X stock, Invitrogen)]) and incubated at 37 °C in 5% CO₂ incubator for 1 day. Wells containing growth medium only were used as background controls. Cells were treated with various concentrations of AFMT in quadruplicate. Two days post treatment, 20 µL of CellTiter 96® AQeuous One Solution Reagent (Promega) was added to each well. The plates were incubated at 37 °C in 5% CO₂ incubator for 2 hours followed by absorbance measurement at 490 nm. To calculate relative cell viability, the readings for each compound concentration were subtracted from the background controls and normalized to vehicle controls.

To investigate AFMT activity in vivo, Jordan Bisanz (Turnbaugh lab, UCSF) administered either AFMT (25 mg/kg) or a vehicle control in combination with L-dopa (10 mg/kg) and carbidopa (30 mg/kg) to gnotobiotic mice colonized with *E. faecalis* MMH594 (Figure 2.39). We found that AFMT significantly increased the peak serum concentration (C_{max}) of L-dopa compared to vehicle (Figure 2.39) (p<0.05, two-tailed Mann Whitney test). While these observations are consistent with our in vitro data and inhibition of first-pass gut microbial metabolism in the intestine, this experiment some important caveats. For example, we cannot rule out the possibility that AFMT directly modulates uncharacterized targets in the mouse. Additionally, it is possible that AFMT is transformed into metabolites that could interfere with host metabolism. For example, AFMT could be hydroxylated by the host tyrosine hydroxylase to form α-fluoromethyldopa, a potent inhibitor of host AADC. This metabolism of AFMT was previously observed in rats but it
is unclear whether it would occur at the relatively short timescales of our experiment (95, 96).
Comparing serum L-dopa levels germ-free mice colonized with either wild-type E. faecalis or a

\textit{tyrDC} mutant would highlight whether microbial L-dopa decarboxylation impacts drug levels in
vivo. If AFMT improves serum L-dopa in both of these groups, one would conclude that the
inhibitor does not act through inhibiting TyrDC. Overall, these data suggest that AFMT could be
a promising tool compound for studying bacterial L-dopa metabolism (97) and highlight the
potential of developing L-dopa-based combination therapies containing drugs that target both host
and gut microbial decarboxylation.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure2_39.png}
\caption{Effect of AFMT on serum L-dopa in gnotobiotic mice colonized with \textit{E. faecalis}
MMH594 and given L-dopa and carbidopa.}
A) Mice treated with AFMT inhibitor demonstrated higher serum L-dopa relative to vehicle
controls. Error bars represent the Mean ± the SEM. B) The maximum serum concentration ($C_{\text{max}}$)
of L-dopa is significantly higher with AFMT relative to vehicle controls. Panels A,B *P<0.05,
mann-Whitney U test; n = 4-5/group. The pharmacokinetic experiment was performed by Jordan
Bisanz (Turnbaugh lab, UCSF). The extraction of blood samples and analytical chemistry was
done by Vayu Maini Rekdal.
\end{figure}
2.3 Discussion

In these initial studies we elucidated the molecular basis of gut microbial L-dopa decarboxylation. Our studies suggest that the prominent human gut microbe *E. faecalis* decarboxylates L-dopa using PLP-dependent decarboxylase. Tyrosine is the preferred substrate of this enzyme, but decarboxylation of L-dopa by purified enzyme, whole cells, and complex communities still proceeds in the presence of physiologically relevant concentrations of tyrosine. This activity promotes growth of the organism under acid stress and is variably distributed among gut microbiotas from Parkinson’s disease patients and neurologically healthy controls. Having uncovered that the widely prescribed inhibitor carbidopa does not inhibit microbial decarboxylation, we developed AFMT as a selective non-toxic inhibitor that blocks microbial decarboxylation in complex microbial community and increases serum L-dopa in mice. Finally, we uncovered organisms and enzymes likely involved in alternative, minor gut microbial pathways that metabolize L-dopa. Altogether, our study deciphers the molecular mechanisms by which gut bacteria could interfere with the treatment of Parkinson’s disease. The decarboxylation of L-dopa by *E. faecalis* mirrors host drug metabolism and, together with human AADC, likely limits drug availability and contributes to interindividual variation in efficacy. Together with recent work dissecting host and gut microbial contributions to the antiviral drug brivudine (98), our findings show that gut bacterial metabolism need not be chemically distinct from host activities to alter drug efficacy and suggest such interactions may be underappreciated. Moreover, carbidopa’s failure to prevent L-dopa decarboxylation by *E. faecalis* implies that additional host-targeted drugs may lack efficacy toward activities also present in the gut microbiota.

Our findings on L-dopa decarboxylation are consistent with a recent study by van Kessel and co-workers (99). Similar to our work, the authors demonstrated that *E. faecalis* and *E. faecium*
decarboxylate L-dopa in anaerobic cultures and that the gut microbe *L. brevis* had only minimal activity despite encoding a TyrDC homolog. Mutagenesis and in vitro biochemistry confirmed TyrDC as responsible for L-dopa decarboxylation in *E. faecalis*, and further characterization of this metabolism by purified enzyme and whole cells revealed that decarboxylation takes place in even in the presence of tyrosine, the preferred substrate of TyrDC. Consistent with our work, they also found that benserazide and carbidopa only weakly inhibited decarboxylation by microbial cells and enzymes. Finally, their study reported that stool tyrDC abundance, as assessed by qPCR, positively correlated with carbidopa/L-dopa dosage and disease duration in Parkinson’s disease patients. The correlation was relatively weak (R² = 0.44 and p = 0.037 for dose; R² = 0.66 and p=0.003 for disease duration) and was only performed in 10 subjects, warranting further work on a larger group of patients. Nonetheless, the authors highlight that the correlation could reflect underlying variability in metabolism, as tyrDC abundance in the rat jejunum (a major site of L-dopa absorption) correlated with plasma L-dopa levels, and conventional rats inoculated with *E. faecalis* WT had significantly lower serum L-dopa than rats inoculated with a tyrDC mutant. Taken together, these results and those of our study suggest that previously overlooked gut microbial L-dopa decarboxylation may play important roles in L-dopa efficacy.

While the findings by van Kessel and co-workers strongly support our findings regarding TyrDC and its interactions with commonly prescribed decarboxylase inhibitors, there are some important differences. For example, they did not demonstrate a potential physiological role for L-dopa decarboxylation in *E. faecalis* or identify organisms and genes involved in alternative, minor pathways for gut microbial L-dopa metabolism or in the dehydroxylation of dopamine into *m*-tyramine. Additionally, although they demonstrate the presence of the tyrDC gene in a small number of complex gut microbiota samples from Parkinson’s patients (n=10), they did not show
that the associated L-dopa decarboxylase activity occurs in these samples. Moreover, despite uncovering the lack of efficacy of carbidopa and benserazide towards microbial L-dopa metabolism, they did not identify strategies for inhibiting the bacterial enzyme. In contrast, we demonstrated that TyrDC predicts direct drug metabolism in Parkinson’s patient microbiotas and use an understanding of its substrate specificity to identify a small molecule that prevents L-dopa decarboxylation in patient samples and increases L-dopa bioavailability in vivo. Through discovering predictive biomarkers for L-dopa metabolism and identifying an inhibitor of this activity, our work will enable efforts to elucidate the contribution of the gut microbiota to drug availability, patient drug response, and treatment outcomes.
2.4 Methods

2.4.1 General materials

The following chemicals were used in this study: L-dopa (Sigma-Aldrich, catalog# D9628-5G), dopamine (Sigma-Aldrich, catalog# PHR1090-1G), Isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich, catalog# I5502-1G), carbidopa (Sigma-Aldrich, catalog# PHR1655-1G), L-tyrosine (Sigma-Aldrich, catalog# T3754-50G), Pyridoxal-L-Phosphate (Sigma-Aldrich, catalog# P9255-1G), SIGMAFAST protease inhibitor tablets (Sigma-Aldrich, catalog#: S8830), (S)-α-fluoromethyltyrosine (AFMT) (obtained from Merck Sharp & Dohme Corp under MTA LKR166502), L-phenylalanine (Sigma-Aldrich, catalog# P2126-100G), δ3-phenyl-L-dopa (Sigma-Aldrich, catalog# 333786-250MG). Luria-Bertani (LB) medium was prepared from its basic components (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) or obtained from either EMD Millipore or Alfa Aesar. Acetonitrile and methanol for LC-MS analyses were purchased as LC-MS grade solvent from Honeywell Burdick & Jackson or Sigma-Aldrich.

2.4.2 General methods

All bacterial culturing work was performed in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen, 10% carbon dioxide, and nitrogen as the balance. Hungate tubes were used for anaerobic culturing unless otherwise noted (Chemglass, catalog# CLS-4209-01). All lysate work and protein experiments were performed in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen and nitrogen as the balance situated in a cold room at 4 °C.

All genomic DNA (gDNA) was extracted from bacterial cultures using the DNeasy UltraClean Microbial Kit (Qiagen, catalog #: 12224-50) according to the manufacturer’s protocol.
All cloning work was performed as follows. NcoI (catalog# R3193S), HindIII (catalog# R0104S), XhoI (catalog# R0146s) and NdeI (R0111L) were purchased from New England Biolabs. For restriction digestion, 500 ng plasmid or 50 ng PCR insert were mixed with 1 µL of each restriction enzyme and 4 µL 10x cutsmart buffer (New England Biolabs, catalog# B7200S), and MilliQ water to a final reaction volume of 40 µL. Restriction digestion reactions were left at 37 °C. for 3 hours, followed by gel purification using the GFX PCR DNA and Gel Band Purification Kit (GE healthcare, catalog# 28-9034-70). Ligation of purified digested vectors and inserts was performed by Gibson Assembly. Briefly, 50 ng of vector was mixed with 3-fold molar excess of insert, 5 µL Gibson Assembly 2x Mastermix (New England Biolabs, catalog# E2611S), and MilliQ water to a final volume of 10 µL. The Gibson reactions were left at 50 °C for 30 minutes and 5 µL of the reaction was transformed into chemically-competent E. coli TOP10 using heat shock.

2.4.3 LC-MS methods

Method A: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire PFP column (4.6 × 100 mm, 3.5 µm; catalog #81601). The flow rate was 1.0 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 0% B isocratic, 2-9 min: 0-10% B, 9-11 min: 10-95% B, 11-13 min: 95% B isocratic, 13-15 min: 95-0% B, 15-18 min: 0% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of $d_3$-phenyl-L-dopa (precursor ion m/z = 201.3, daughter ion m/z = 155.2), L-dopa (precursor ion m/z = 198.3, daughter ion m/z = 152.2), $d_3$-phenyl-dopamine (precursor ion m/z = 157.3, daughter ion m/z = 140.3), dopamine (precursor ion m/z = 154.3, daughter ion m/z = 137.3), $d_3$-phenyl-tyramine (precursor
ion m/z = 141.3, daughter ion m/z = 124.3), tyramine (precursor ion m/z = 138.3, daughter ion m/z = 121.3) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode.

Method B: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Phenomenex Kinetex 5 µm Biphenyl 100 Å (50*4.6 mm, product #: 00B-4627-E0). The flow rate was 0.4 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-6 min: 0% B isocratic. The same masses as in Method A were monitored using the same settings.

Method C: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim polar advantage II (3 µM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. The same masses as in Method A were monitored using the same settings.

Method D: Identical to Method A, but an additional mass of phenylethylamine (precursor ion m/z = 122.3, daughter ion m/z = 105.3) was monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode.

Method E: Samples were analyzed using an Agilent technologies 6530 Accurate-Mass Q-TOF LC/MS and a Phenomenex Kinetex 5 µm Biphenyl 100 Å (50*4.6 mm, product #: 00B-4627-E0). The flow rate was 0.4 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 5% B isocratic, 2-25 min: 0-95% B,
25-30 min: 95% B isocratic, 30-40 min: 95-5% B. For the MS detection, the ESI mass spectra data were recorded on a positive mode for a mass range of m/z 50 to 3000. A mass window of ± 0.005 Da was used to extract the ion of [M+H].

Method F: Identical to Method A, but an additional mass of tyrosine (precursor ion m/z = 182.3, daughter ion m/z = 136.2) was monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode.

Method G: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim polar advantage II (3 μM, 120 Å, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-8 min: 99% C isocratic, 8-9 min: 99-50% B, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of $d_3$-phenyl-dihydrocaffeic acid (precursor ion m/z = 184.2, daughter ion m/z = 140.2), $d_3$-phenyl-dihydroxyphenylacetic acid (precursor ion m/z = 170.2, daughter ion m/z = 126.2), $d_3$-phenyl-hydroxyphenylpropionic acid (precursor ion m/z = 168.2, daughter ion m/z = 124.2), $d_3$-phenyl-hydroxyphenylacetic acid (precursor ion m/z = 154.2, daughter ion m/z = 110.2) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode.

Method H: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801). The flow rate was 0.5 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-9 min: 0-10% B, 9-11 min: 10-95% B, 11-13 min: 95% B isocratic, 13-14 min: 95-0% B, 14-18 min: 0% B isocratic. For mass spectrometry,
the source temperature was 300 °C, and the masses of $d_3$-phenyl-L-dopa (precursor ion m/z = 201.3, daughter ion m/z = 155.2), L-dopa (precursor ion m/z = 198.3, daughter ion m/z = 152.2), $d_3$-phenyl-dopamine (precursor ion m/z = 157.3, daughter ion m/z = 140.3), dopamine (precursor ion m/z = 154.3, daughter ion m/z = 137.3), $d_3$-phenyl-tyramine (precursor ion m/z = 141.3, daughter ion m/z = 124.3), tyramine (precursor ion m/z = 138.3, daughter ion m/z = 121.3) were monitored at a collision energy of 15mV and fragmentor setting of 135 in MRM mode.

2.4.4 BLASTP search for *L. brevis* TyrDC in Human Microbiome Project reference isolate genomes

This analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF). The *L. brevis* tyrosine decarboxylase (UniProt accession, B8V35) was used as the query sequence for a BLASTP search of the Human Microbiome Project (HMP) reference isolate genomes. All GenBank assemblies associated with the Human Microbiome Project Reference Genome project (PRJNA28331) were retrieved (20 June 2018). Protein sequences were collected were queried by BLASTP (BLAST+ 2.6.0) requiring a minimum e-value of 1e-4. Separately, scaffolds were queried for 16S rRNA genes (github.com/tseemann/barrnap). Where multiple 16S rRNA copies were present, the longest was selected, and if multiple equal length copies were present, the first was selected. An alignment was carried out using muscle (100) with the following parameters: `-maxiters 3 -diags1 -sv`. The alignment was trimmed removing positions where >20% of alignment was gapped (101) and a tree was built using FastTree (102). The tree was subsequently rooted on *Methanobrevibacterial* species and represented as a cladogram with ggtree (103).

2.4.5 Assessment of conservation of TyrDC across *E. faecalis* genomes

This analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF). The amino acid sequence of TyrDC from *E. faecalis* TX0645 (UniProtKB #E6I994) was subjected to a tBLASTn
search against whole genome shotgun contigs in NCBI using the default BLAST parameters. All but 3 of the 655 *E. faecalis* genomes that were searched had the TyrDC sequence conserved at 98% amino acid ID and 100% query coverage.

### 2.4.6 Screen for L-dopa decarboxylation in anaerobic bacterial cultures

The strains screened for L-dopa decarboxylation were *Enterococcus raffinosus* ATCC 49464, *Enterococcus mundtii* ATCC 882, *Enterococcus villorum* LMG 12287T, *Enterococcus dispar* ATCC 51266, *Enterococcus massilensis* DSM 100308, *Carnobacterium maltaromaticum* ATCC 35586, *Enterococcus moraviensis* ATCC BAA-383, *Enterococcus hirae* ATCC 8043, *Enterococcus haemiperoxidus* ATCC BAA-382, *Enterococcus durans* ATCC 6056, *Enterococcus caccae* ATCC BAA-1240, *Enterococcus faecalis* MMH549, *Enterococcus faecalis* V583, *Enterococcus faecalis* OG1RF, *Enterococcus faecalis* TX0104, *Enterococcus faecium* E1007, *Enterococcus faecium* E2134, *Enterococcus faecium* TX01330, *Providencia rettgeri* DSM 1131, *Proteus mirabilis* ATCC 29906, and *Lactobacillus brevis* subsp. *gravesensis* ATCC 27305. All *Enterococcus* and *Carnobacterium* strains were grown in Brain Heart Infusion (BHI) broth (Beckton Dickinson, catalog# 211060), while *P. rettgeri* and *P. mirabilis* were grown in MEGA medium (104). *L. brevis* was grown in MRS medium (Sigma-Aldrich, catalog# 69966-500G). Starter cultures were grown from individual colonies for 48 hours at 37 °C in liquid media and were then inoculated 1:100 into triplicate tubes containing fresh media and 500 µM L-dopa. Cultures were grown for 48 hours at 37 °C, after which they were harvested by centrifugation. Culture supernatant was diluted 1:20 in LC-MS grade MeOH, and 5 µL of this supernatant was analyzed on LC-MS/MS using Method A described above. % Decarboxylation was calculated as the concentration of dopamine divided by the total concentration of dopamine and L-dopa in the culture. All experiments were performed in an anaerobic chamber.
2.4.7 Confirmation of *E. faecalis* MMH594 tyrDC mutants

Mutants were generated as part of a transposon mutagenesis library and were generously provided by the Gilmore lab at Massachusetts General Hospital. The presence of the 2 kb Tet cassettes were verified by PCR using primer 1 (ATGAAAAACGAAAAATTAGCAAAAAGG) and primer 2 (CATACCAAGGCTTCTAAGTTCAGC) for mutants 1 and 2 and primer 3 (CATTCAAGAGAAGTGTATGATGC) and primer 4 (CAGTATGGAAGTAAATGATCCG) for mutants 3 and 4. The genomic location of the mutants is shown below in table 2.4.

### Table 2.4 Genomic locations of transposon mutants in *E. faecalis* MMH594.

<table>
<thead>
<tr>
<th>Mutant name</th>
<th>Position of Tn in genome</th>
<th>Locus Tag</th>
<th>Primer pairs used to confirm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔtyrDC_1</td>
<td>568707</td>
<td>SAW_00626T0</td>
<td>1,2</td>
</tr>
<tr>
<td>ΔtyrDC_2</td>
<td>568897</td>
<td>SAW_00626T0</td>
<td>1,2</td>
</tr>
<tr>
<td>ΔtyrDC_3</td>
<td>570186</td>
<td>SAW_00626T0</td>
<td>3,4</td>
</tr>
<tr>
<td>ΔtyrDC_4</td>
<td>569734</td>
<td>SAW_00626T0</td>
<td>3,4</td>
</tr>
</tbody>
</table>

2.4.8 Cloning and expression of *E. faecalis* MMH594 tyrosine decarboxylase

*tyrDC* was amplified from *E. faecalis* MMH594 genomic DNA with primers 5’ TGGTGCCGCGCAGCCATATGAAAAACGAAAAATTAGCACAAGG 3’ and 5’ TGGTGGTGGTGCTGAGTTTTTACGTCGAAATTTGTTC 3’. The purified PCR product and the pET-28a vector were digested with XhoI and NdeI and, following purification of the digested vector, were ligated together using Gibson assembly. The final vector pET28a encoding N-His$_6$ TyrDC was confirmed to contain the appropriate insert by Sanger sequencing. For protein expression, the pET-28a N-His$_6$ TyrDC vector was transformed into chemically competent *E. coli* BL21(DE3). Starter cultures were grown from individual colonies in 5 mL of LB medium containing 50 µg/mL kanamycin for 18 hours and were then inoculated into 500 mL
of LB medium containing 50 µg/mL kanamycin. Cells were grown with shaking at 37 °C until reaching OD₆₀₀=0.400, at which point isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 200 µM. Cells were grown for 18 hours at 15 °C with shaking. For purification of recombinant TyrDC, the 500 mL culture was harvested by centrifugation, and the resulting pellet was resuspended in 40 mL 50 mM Tris pH containing 0.25M NaCl, followed by lysis using a cell disruptor (Avestin Emulsiflex C3). All of the clarified lysate was loaded onto 2 mL of HisPur Ni-NTA resin (Thermo Fisher Scientific, Waltham, MA) and eluted using a gradient of 50 mM to 200 mM imidazole (in 50 mM Tris pH 8 containing 0.25M NaCl). Fractions containing pure protein were combined and dialyzed over two rounds into in 50 mM Tris pH 8 containing 0.20 M NaCl and 10% w/v glycerol. The dialyzed protein was concentrated 10-fold using spin columns (VMR, catalog # 97027-9) and frozen at a final concentration of ~100 µM in liquid nitrogen. Protein aliquots were stored at −80 ºC.

2.4.9 Michaelis–Menten kinetics of *E. faecalis* TyrDC

To assess the steady-state kinetics of TyrDC towards tyrosine and L-dopa, the enzyme was first pre-incubated with pyridoxal-5’-phosphate (PLP) at a ratio of 1:1333 in 0.2 M sodium acetate buffer, pH 5.5 for five minutes. The enzyme reaction was then initiated by dissolving the preincubated enzyme-PLP mix 1:10 with substrate dissolved in 0.2 M sodium acetate buffer, pH 5.5. The final concentrations in the final enzyme reaction were 200 µM PLP, 0.15 µM enzyme and 100 µM-1800 µM tyrosine or 100 µM-2250 µM L-dopa. Aliquots of the enzyme reaction were quenched by diluting 10-fold into ice cold methanol at 20, 40, and 60 seconds after initiation. The quenched reaction was centrifuged to pellet precipitates and 5 µL of the supernatant was analyzed by LC-MS/MS using Method B described above. Rates were calculated as the substrate produced from 20 to 60 seconds and were fit to a standard Michaelis–Menten kinetics curve in GraphPad.
Prism (version 7). All experiments were done in triplicate and were repeated at least twice. Reactions were performed at room temperature.

2.4.10 Competition of L-dopa and tyrosine for reaction with purified *E. faecalis* TyrDC

The buffers described above were used for each enzyme in this assay. 0.15 µM TyrDC was mixed with 200 µM PLP and equimolar concentrations of L-dopa and tyrosine (both at a final concentration of 500 µM). Formation of the corresponding amine products was measured by LC-MS/MS following quenching with ice cold methanol (1:10) at specific time points. The quenched mixture was then centrifuged for 10 minutes to pellet precipitates. Method D was used for LC-MS/MS analysis of the supernatant as described above. Reactions were performed at room temperature.

2.4.11 Assessment of the impact of pH on L-dopa and tyrosine decarboxylation by *E. faecalis*

*E. faecalis* MMH594, *E. faecalis* V583, *E. faecalis* TX0104, *E. faecalis* OG1RF were inoculated from single colonies into 10 mL of BHI medium in individual Hungate tubes. Following 24 hours of anaerobic growth, these starter cultures were inoculated in triplicate 1:10 into 200 µL of BHI medium (pH 5 or pH 7) and were grown at 37 ºC anaerobically in 96-well plates (VWR, catalog# 29442-054). Plates were set up in duplicate. One of these plates was used to measured growth in the Synergy HTX Multi-Mode Microplate Reader (BioTek) by measuring absorbance at 600 nm. The other plate was used for withdrawing culture aliquots for metabolite analysis; at 0, 2, 4, 8, 24 hours, a 30 µL aliquot was removed from each culture and immediately frozen at -20 ºC for downstream metabolite analysis. Thawed aliquots were centrifuged to pellet the cells and supernatants were diluted 1:10 in LC-MS grade methanol and were analyzed by LC-MS using Method F described above.
2.4.12 Assessment of the impact of pH and tyrosine on L-dopa decarboxylation by *E. faecalis* MMH594

BHI medium was adjusted to the appropriate pH with HCl prior to autoclaving. *E. faecalis* MMH594 was inoculated from single colonies into 10 mL of BHI medium in individual Hungate tubes. Following 24 hours of anaerobic growth, these starter cultures were inoculated in triplicate 1:10 into 200 µL of BHI medium (pH 5 or pH 7) containing either 0 or 1 mM L-tyrosine added. The final tyrosine concentrations in the BHI medium were approximately 500 µM without added tyrosine and 1500 µM with tyrosine added. These two concentrations approximate the resting and post-meal small intestinal tyrosine concentration in healthy human volunteers (40). The cultures were grown at 37 ºC anaerobically in 96-well plates (VWR, catalog# 29442-054). Plates were set up in duplicate. One of these plates was used to measured growth in the Synergy HTX Multi-Mode Microplate Reader (BioTek) by measuring absorbance at 600 nm. The other plate was used for withdrawing culture aliquots for metabolite analysis; at 0, 2, 4, 8, 24 hours, a 30 µL aliquot was removed from each culture and immediately frozen at −20 ºC for downstream metabolite analysis. Thawed aliquots were centrifuged to pellet the cells and supernatants were diluted 1:10 in LC-MS grade methanol and were analyzed by LC-MS using Method F described above.

2.4.13 Assessment of the impact of pH and tyrosine on L-dopa decarboxylation by three fecal samples from neurologically healthy humans

BHI medium was adjusted to the appropriate pH using HCl prior to autoclaving. Fecal slurries had been previously frozen in PBS with 20% glycerol. These samples were thawed at room temperature anaerobically and inoculated 1:200 into 5 mL BHI medium (pH 5 or pH 7) with either 0 or 1 mM L-tyrosine added. The final tyrosine concentrations in the BHI medium were approximately 500 µM without added tyrosine and 1500 µM with tyrosine added. These two
concentrations approximate the resting and post-meal small intestinal tyrosine concentration in healthy human volunteers (68). Experiments were performed anaerobically, and cultures were grown in Hungate tubes at 37 ºC. At 0, 11, 23, 32, and 48 hours, a 200 µL aliquot was removed from each culture. Following measurement of absorbance at 600 nm using a Synergy HTX Multi-Mode the Microplate Reader (BioTek), aliquots were immediately frozen at –20 ºC for downstream metabolite analysis. The thawed aliquots were centrifuged to pellet the cells, and supernatants were diluted 1:10 in LC-MS grade methanol and were analyzed by LC-MS using Method F described above.

2.4.14 Calculation of drug concentrations in the small intestine

The maximum molar concentration (molarity) likely to be achieved by a given drug was calculated by converting the dose amount into moles and then dividing by 100 mL, the approximate volume of the small intestine after drinking 240 mL of water (105). For levodopa, the dose range considered was 0.10 to 6.0 grams (33), which resulted in a range of 0.51 to 30.4 mmol and thus a concentration range of 5.1 to 304 mM. For carbidopa, the dose range considered was 10 mg to 200 mg (33), which resulted in a range of 0.044 to 0.88 mmol and thus a concentration range of 0.44 to 88 mM.

2.4.15 Cloning and heterologous expression of H. sapiens aromatic amino acid decarboxylase

H. sapiens aromatic amino acid decarboxylase (AADC) was obtained as cDNA from Sino Biological (catalog#: HG10560-M). The cDNA was amplified using primers 5’ AACCATGGATGAACGCAAGTGAATTCCGAAGG 3’ and 5’ GGAAGCTTCTCCCTCCTCGTCGACGC 3’. PCR amplicons were purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE healthcare, catalog# 28-9034-70) The purified PCR amplicon as well as pET28a were digested by NeoI and HindIII restriction enzymes using
standard protocols. The insert was ligated into pET-28a using Gibson assembly. The final pET-28a vector encoding C-His\textsubscript{6} AADC was confirmed to contain the appropriate insert by Sanger sequencing. For protein expression, the pET-28a C-His\textsubscript{6} AADC vector was transformed into chemically competent \textit{E. coli} BL21 (DE3). Starter cultures were grown from individual colonies in 5 mL of LB medium containing 50 µg/mL kanamycin for 18 hours and were then inoculated into 4 L of LB medium containing 50 µg/mL kanamycin. When OD\textsubscript{600} reached 0.500, IPTG was added at a final concentration of 200 µM to induce protein expression. Cells were grown for 18 hours at 15 °C with shaking. The purification of AADC followed the general protocol outlined by Montioli \textit{et al.} (106) with minor modifications. Briefly, cultures were harvested by centrifugation and the resulting pellet was resuspended in 40 mL 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M NaCl, 20 mM imidazole, 4 mg/mL protease inhibitor tablets (Sigma-Aldrich, catalog#: S8830) and 50 µM PLP, followed by lysis using a cell disruptor (Avestine emulsiflex C3). All the clarified lysate was loaded onto Ni resin and eluted in a gradient from 50 mM to 200 mM imidazole (in 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M NaCl). Pure fractions were combined and dialyzed over two rounds into in 100 mM sodium phosphate buffer pH 7.4 and 10% w/v glycerol. The dialyzed protein was concentrated 10-fold using spin columns and frozen at a final concentration of ~80 µM in liquid nitrogen. Protein aliquots were stored at −80 °C.

2.4.16 Evaluation of inhibitors towards \textit{E. faecalis} and \textit{H. sapiens} decarboxylases

To assess the inhibitory effects of carbidopa on L-dopa decarboxylation, the general process and setup described in (107) was followed, with some modifications. For TyrDC, enzyme was first pre-incubated at a ratio of 1:1333 with PLP in 0.2 M sodium acetate buffer, pH 5.5 for five minutes. The enzyme reaction was then initiated by dissolving the preincubated enzyme-PLP mix 1:10 with substrate and inhibitor dissolved in 0.2 M sodium acetate buffer, pH 5.5. The final concentrations
in the enzyme reaction were 200 µM PLP, 0.15 µM enzyme, 500 µM L-dopa and 0.5 µM-1000 µM carbidopa. Aliquots of the enzyme reaction were quenched by diluting 10-fold into methanol at 20, 40, and 60 seconds after initiation. For AADC, enzyme was pre-incubated for five minutes at a ratio of 1:500 with PLP in 20 mM sodium phosphate buffer, pH 7.4. The enzyme reaction was then initiated by dissolving the preincubated enzyme-PLP mix 1:10 with substrate and inhibitor dissolved in 20 mM sodium phosphate buffer, pH 7.4. The final concentrations in the enzyme reaction were 100 µM PLP, 0.2 µM enzyme, 500 µM L-dopa and 0.01 µM-10 µM of carbidopa. Aliquots of the enzyme reaction were quenched by diluting 10-fold into methanol at 60, 120, and 180 seconds after initiation. For (S)-α-fluoromethyltyrosine (AFMT), the same buffer conditions as above were used for each enzyme but following a slightly different setup. For TyrDC, 1 µM enzyme was first preincubated with 2000 µM PLP and AFMT (0.5-20 µM) for 14 minutes while for AADC, 1 µM enzyme was first preincubated with 200 µM PLP and AFMT (10-650 µM), for 14 minutes. Reactions were initiated by diluting this mixture 1:10 in buffer containing 500 µM L-dopa. The TyrDC reaction was quenched by diluting 1:10 in MeOH after 60 seconds while the AADC reaction was quenched after 120 seconds. For both AADC and TyrDC, the quenched reactions were spun down, and 5 µL of the supernatant was analyzed by LC-MS/MS using Method B described above. IC50 curves were fit in GraphPad Prism (version 7). All experiments were done in triplicate and were repeated at least twice. Rates were calculated as the substrate produced over the timepoints collected and were normalized to the rate without inhibitor to produce a measure of % activity. Reactions were performed at room temperature.
2.4.17 Identification of a PLP-inhibitor adduct upon incubation of enzymes with AFMT

The same buffers as described above were used for each enzyme. 10 µM AADC or TyrDC was incubated with 200 µM PLP and 50 µM AFMT for one hour at room temperature. The reactions were quenched by boiling at 95 ºC for 15 minutes. This mix was diluted 1:20 in LC-MS grade methanol and spun down to pellet precipitates. The supernatant was analyzed by LC-MS using Method E described above. The predicted adduct was identified based on the work described in (91).

2.4.18 Substrate screen with *E. faecalis* TyrDC and *H. sapiens* AADC

The same buffers as described above were used for each enzyme. 0.15 µM TyrDC was mixed with 200 µM PLP, and 0.2 µM AADC was mixed with 200 µM PLP. L-dopa, phenylalanine, m-tyrosine, or p-tyrosine was added to each of the enzymes at a final concentration of 500 µM. The AADC reaction was quenched at 180 seconds and the TyrDC reaction was quenched at 60 seconds, both by diluting 1:10 in ice cold methanol. The plates were centrifuged to pellet precipitates and supernatants were analyzed on LC-MS/MS using Method D described above. Reactions were performed at room temperature.

2.4.19 Evaluation of inhibitors towards growth and L-dopa decarboxylation by *E. faecalis*

*E. faecalis* MMH594 starter cultures were grown in BHI medium anaerobically over 12 hours from individual colonies at 37 ºC. The starter culture was diluted 1:10 in 180 µL of fresh BHI medium containing 500 µM L-dopa and varying concentrations of carbidopa or AFMT. Cultures were grown anaerobically at 37 ºC for 18 hours and harvested by centrifugation. Supernatants were diluted 1:10 in LC-MS grade methanol and were analyzed by LC-MS using Method C described above. % Inhibition was calculated as the concentration of L-dopa remaining relative to that remaining in culture supernatants from the *E. faecalis* MMH594 *tyrDC* mutant grown under
the same conditions. Growth was monitored using a Synergy HTX Multi-Mode Microplate Reader (BioTek) by measuring absorbance at 600 nm. Experiments were performed anaerobically, and cultures were grown in 96-well plates (VWR, catalog# 29442-054).

2.4.20 Collection of fecal samples from neurologically healthy human patients

The indicated stool samples were collected from 12 neurologically healthy individuals during the control phase of an inpatient study described in detail elsewhere [108] and from 7 healthy control subjects sampled at the University of California, San Francisco (UCSF). All subjects consented to participate in the study, which was approved by the relevant Institutional Review Boards. These were collected by the Turnbaugh lab at UCSF. Samples were stored as fecal slurries in PBS and 20% glycerol at –80 °C until use.

2.4.21 Collection of fecal samples from Parkinson’s disease patients

The indicated stool samples were obtained from the BioCollective Microbiome Stool Bank (Denver, CO), which collected the samples under the Western IRB WIRB #1164096 for human subjects research. Samples were stored as fecal slurries in PBS and 20% glycerol at –80 °C until use. Specific samples (n = 12) were selected from a larger set to minimize confounding variables, and they include 6 drug-naïve individuals and 6 patients taking L-dopa/carbidopa. Participants were not taking: antibiotics, antihistamines, laxatives, suppositories, beta blockers, statins, proton pump inhibitors, tricyclic antidepressents, selective serotonin reuptake inhibitors (SSRIs), platelet aggregators, oral contraceptives, oral metformin, and nonsteroidal anti-inflammatory drugs. They did not have additional comorbidities. Groups were balanced on gender (drug-naïve = 3 males and 3 females vs. L-dopa/carbidopa = 4 males and 2 females), age (drug-naïve = 68 ± 6 years vs. L-dopa/carbidopa = 60 ± 8 years), and body mass index (drug-naïve = 23.1±4.4 kg/m² vs L-
dopa/carbidopa = 25.3±3.8 kg/m²) for drug naïve and L-dopa/carbidopa patients respectively (mean ± standard deviation (SD)).

2.4.22 Ex vivo assays with \(d_3\)-phenyl-L-dopa

First, we prepared fecal slurries: stool samples were resuspended in pre-reduced PBS at a final concentration of 0.1 g/mL. The mixture was vortexed to produce a homogenous slurry and was then spun down at 600g for 10 minutes. It was then left for 30 minutes to let particulates settle. These slurries were diluted to a final volume of 300 µL in 20% glycerol and then aliquoted into individual tubes, flash-frozen in liquid nitrogen and stored at –80 ºC, allowing for experiments to be repeated with the same fecal sample multiple times without further freeze-thawing. These slurries were diluted 1:100 into 5 mL MEGA or BHI media containing 1 mM \(d_3\)-phenyl-L-dopa with or without carbidopa (2 mM) or AFMT (200-250 µM) in triplicate and were grown anaerobically at 37 ºC for 72 hours. Cultures were harvested by centrifugation and culture supernatants were diluted 1:10 in LC-MS grade methanol, followed by another round of centrifugation to pellet precipitates. Supernatants were analyzed by LC-MS using Method C or Method G described above.

2.4.23 qPCR assays

gDNA was extracted from the culture pellets generated in the experiments described above (‘Ex vivo assays with \(d_3\)-phenyl-L-dopa’) using the DNeasy UltraClean Microbial Kit. 2 ng of the extracted DNA from each culture was used for qPCR assays containing 10 µL of iTaq Universal SYBRgreen Supermix (Bio-rad, catalog 3: 1725121), 7 µL of water, and 10 µM each of forward and reverse primers. PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 95 ºC for 5 minutes 34 cycles of 95 ºC for 1 min, 60 ºC for 1 min, 72 ºC for 1 min. The program ended with a final extension at 34 ºC for 5 mins. The
primers used were: primers for *E. faecalis* 16S rRNA (109): 5’ CGCTTCTTTTCTCCCGAGT 3’ and 5’ GCCATGCGGCATAAACTG 3’; primers for *tyrDC* (110): 5’ CGTACACATTTCATGCTGATGCGAT 3’ and 5’ATGTCCTACTCCTCCTCCCATTTG 3’.

**2.4.24 Metagenomic analysis of *E. faecalis*/*tyrDC* abundance and prevalence across human patients**

This analysis was performed by Dr Jordan Bisanz (Turnbaugh lab, UCSF). A curated collection of human gut microbiomes representing 1870 individuals (111), was used to correlate the abundances of *Enterococcus*/*tyrDC* (Pearson correlation, R). Prevalence was estimated as a function of rolling minimum abundance cut off.

**2.4.25 MTT assay for HeLa cell viability in the presence of AFMT**

HeLa cells were seeded into 96-well plates at a density of 1 × 10⁵ cells per well in 100 μL of growth medium [(DMEM medium supplemented with 10% FBS (2 mL) and 1X Antibiotic-Antimycotic (100X stock, Invitrogen))] and incubated at 37 °C in a 5% CO₂ incubator for 1 day. Wells containing growth medium only were used as background controls. Cells were treated with various concentrations of AFMT in quadruplicate. Two days post treatment, 20 μL of CellTiter 96® AQeuous One Solution Reagent (Promega) was added to each well. The plates were incubated at 37 °C in a 5% CO₂ incubator for 2 hours followed by measurement of absorbance at 490 nm using a Synergy HTX Multi-Mode Microplate Reader (BioTek). To calculate relative cell viability, the readings for each compound concentration were subtracted from the background controls and normalized to vehicle controls.

**2.4.26 Pharmacokinetic experiment in gnotobiotic mice**

A preliminary experiment was performed by Dr. Elizabeth Bess when she was a postdoctoral fellow in the Turnbaugh lab at UCSF. The final experiment presented in this work
was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF). Germ-free male BALB/c mice aged 7-12 weeks (n=11) were colonized with $1 \times 10^8$ CFU *E. faecalis* MMH594 (in 100 µL Brain Heart Infusion media) and randomly selected to receive 10 mg/kg AFMT or vehicle control (0.25% carboxymethylcellulose) balancing age across groups. 5 mice received the inhibitor and 6 mice received the control. Mice were administered AFMT or vehicle 1 hour before a second dose co-administered with 10 mg/kg $d_3$-phenyl-L-dopa and 30 mg/kg carbidopa. Blood was collected from the tail vein at 0 min, 15 min, 30 min, 60 min, 90 min, and 120 min for determination of serum L-dopa concentration.

**2.4.27 Extraction and analysis of L-dopa from tail vein blood**

All manipulations were performed in 1.5 mL Eppendorf tubes. 5 µL of tail-vein blood was added to a mixture of 25 µL of 0.2 M sodium acetate buffer (pH 5.5, containing 12.5 µM unlabeled L-dopa as the internal standard) and 50 µL of ice-cold methanol (containing 35 µM EDTA and 0.1 % (w/v) ascorbic acid as anti-oxidants). The mixture was vortexed and left on ice for 5 minutes to precipitate proteins, whereby 100 µL of chloroform was added to remove serum lipids. The mixture was vortexed until the solution appeared homogenous and samples were then left at −80 °C for 10 minutes. Organic and aqueous layers were then separated using centrifugation for 10 minutes, and the top aqueous layer was then carefully removed (50 µL) and transferred to a new Eppendorf tube. This aqueous layer was evaporated using a Genevac (EZ-2 Elite personal evaporator, HPLC setting with 30 mins each for the first and second phases, no heating). The dried residue was then resuspended in 15 µL MilliQ water (containing 0.1 % formic acid) by vortexing. Samples were spun down to pellet potential particulates and 10 µL of the resulting supernatant was injected for LC-MS analysis using Method H described above. Quantification was performed by normalization of the $d_3$-phenyl-L-dopa analyte peak area to the peak area of the unlabeled L-
dopa internal standard.

**2.4.28 Incubation of wild-type and mutant *C. sporogenes* strains with d₃-phenyl-L-dopa.**

Mutants were generated by the Sonnenburg laboratory at Stanford University as part of a previous study (77). Experiments were performed in Hungate tubes and strains were growth anaerobically in TYG medium. The TYG medium was by prepared by adding 3 g tryptone (Beckton Dickinson, catalog# BD 211705), 3 g yeast extract, 3 g glucose, and 1 g potassium phosphate dibasic to 1 L of water followed by adjustment of pH to 7.4 and autoclaving for 15 minutes. For the incubations, *C. sporogenes* ATCC 15579 was first inoculated from a single colony into 10 mL of TYG medium and grown for 48 hours to generate turbid starter cultures. These starter cultures were then diluted into TYG medium containing 1 mM d₃-phenyl-L-dopa in triplicate and were grown anaerobically at 37 °C for 72 hours. Cultures were harvested by centrifugation and culture supernatants were diluted 1:10 in LC-MS grade methanol, followed by another round of centrifugation to pellet precipitates. Supernatants were analyzed by LC-MS using Method G described above.
2.5 References


61. T. Koyanagi, A. Nakagawa, H. Sakurama, K. Yamamoto, N. Sakurai, Y. Takagi, H. Minami, T. Katayama, H. Kumagai, Eukaryotic-type aromatic amino acid decarboxylase from the root
colonizer Pseudomonas putida is highly specific for 3,4-dihydroxyphenyl-L-alanine, an allelochemical in the rhizosphere. Microbiol-Sgm 158, 2965-2974 (2012).


Chapter 3. Discovery of gut microbial strains and enzymes that dehydroxylate dopamine

This chapter was adapted in part from:

3.1 Introduction

Having uncovered gut microbial strains, genes, and enzymes involved in converting L-dopa to dopamine, we next sought to understand the molecular basis of dopamine dehydroxylation. Although it has received less attention than other major neurotransmitters found in the human gut, such as serotonin, dopamine is an important player in gut physiology that exerts its effects through both neuronal and non-neuronal processes (Figure 3.1). The five main subtypes of dopamine receptors (D₁-D₅) are present in the gastrointestinal tract (I). These receptors are expressed during early development of the enteric nervous system and are found both along the length of the gut and from the lumen to mucosa and submyenteric plexus in adults (I, 2). Importantly, enteric neurons express one or more receptor subtypes depending on their location, making the physiological effects of dopamine highly context-dependent (I). For example, dopamine inhibits gastrointestinal motility in the upper gut but stimulates motility in the lower gut (I, 3-5). While the precise receptors mediating these effects are still unclear, knock-out studies in mice suggest that inhibitory signals act through the D₂ dopamine receptor subtype (I).

In addition to its roles in the enteric nervous system, dopamine is produced and sensed by diverse non-neuronal cells, including immune cells, parietal cells, and epithelial cells (6-8). Beyond the enteric nervous system, dopamine promotes colonic mucus secretion, modulates sodium uptake, stimulates gastric acid secretion, contributes to iron homeostasis in macrophages, facilitates water absorption, and promotes antibody production (7, 9-13) (Figure 3.1). The gastrointestinal tract produces over 40% of the dopamine found in the body, suggesting that gut-derived dopamine may contribute significantly to the overall dopaminergic system in the body (6). While some foods such as bananas contain free dopamine, diet is only a minor source of this neurotransmitter in the gut environment (6, 14).
Dopamine is available to gut-dwelling bacteria. This neurotransmitter, along with the related catecholamine neurotransmitter norepinephrine, has been detected in the intestinal lumen of conventional mice (Figure 3.1) (13). Dopamine concentrations vary along the length of the colon, ranging from 16 ng dopamine per g of luminal contents in the ileum, to 93 ng/g in the cecum and 132 ng/g in the colon, the site of the highest microbial density (13). In contrast, germ-free mice harbor barely detectable levels of free dopamine, with a vast majority of this neurotransmitter being present as a host-derived glucuronide conjugate. Colonization of germ-free mice with a mixture of 46 different Clostridia or with stool samples from conventional animals depleted the levels of conjugated dopamine while increasing levels of free dopamine (13). These data suggest that gut bacteria release free catecholamines from host-derived glucuronide conjugates secreted into the gut lumen (Figure 3.1). While the precise microbial enzymes responsible for the dopamine glucuronide hydrolysis have not yet been defined, the UDP-glucuronosyltransferase (UGT) UGT1A10, which is expressed in the small intestine and the colon in humans, is likely responsible for dopamine glucuronidation in the host (Figure 3.1) (15).

Intestinal pathogens utilize free catecholamines for growth (Figure 3.1). Dopamine and norepinephrine promote growth of diverse intestinal pathogens, including *Escherichia coli*, *Yersinia enterolytica*, and *Salmonella enterica* in vitro, and have recently been demonstrated to promote virulence in mice (Figure 3.1) (16-22). For example, mice infected with *S. enterica* Serovar Typhimurium and injected intraperitonially with 5 mg/kg dopamine exhibit significantly reduced survival, dying from infection almost 12 hours earlier than vehicle-treated controls (21). The underlying mechanism involves binding of the intact catecholamine to the bacterial histidine kinase QseC, which activates a range of virulence genes (21, 23). In addition to this mode of transcriptional regulation, dopamine and the related catecholamines epinephrine and
norepinephrine facilitate import of iron into bacterial cells through a number of processes, including direct chelation by the catechol group and increasing expression of iron-uptake genes (21, 22, 24). This helps microbial pathogens overcome the severe iron limitation imposed by the host immune system during infection, contributing to overall pathogenesis (19, 20, 25, 26). Because both dopamine and norepinephrine are produced at substantial levels in the gastrointestinal tract, catecholamine responses in intestinal pathogens may have evolved to indicate their entry into the human gut (6, 27, 28).

Gut microbial metabolism represents another type of microbe-catecholamine interaction. As described in Section 2.1, dopamine dehydroxylation was first discovered as the second step in the major pathway for gut microbial L-dopa metabolism (Figure 3.1) (29, 30). This pathway could contribute play a role in the side effects of L-dopa, a suggestion initially made by Sandler and co-workers in 1974 (30). The authors compared the efficacy and side effects of a standard dose of orally ingested L-dopa and an enteric coated formulation that would release the drug slowly in the gastrointestinal tract. While the two formulations produced no difference in the neurological benefit, after a slight lag period, 75% of patients receiving the enteric coated formulation reported such severe nausea that they had to interrupt the study. This side effect was accompanied by increases in urinary excretion of microbially derived m-tyramine and its host-oxidized product m-hydroxyphenylacetic acid (m-hpaa) (Figure 3.1), with levels of m-hpaa increasing from 14.4 mg to 59.4 mg on average. These data implicated gut microbial dopamine dehydroxylation in L-dopa side effects. Similar to enteric coated L-dopa, small intestinal bacterial overgrowth (SIBO) – a gastrointestinal condition that affects more than 50% of Parkinson’s patients – is associated with increased dopamine dehydroxylation (31, 32). SIBO-positive patients display decreased L-dopa efficacy and higher motor dysfunction, raising the possibility that the increased microbial
metabolism could play a role (33, 34). However, it is also possible that microbial metabolism could decrease the side effects of peripherally produced dopamine, which on its own causes gastrointestinal side effects, and can also result in orthostatic hypotension through activation of vascular dopamine receptors, and cardiac arrhythmias (35, 36). Finally, \textit{m}-tyramine has been speculated to increase L-dopa efficacy, but experimental support for this idea is lacking (37).

The mechanisms underlying the potential host effects of microbially derived L-dopa metabolites are obscure. This gap is partially because the biological targets and functions of \textit{m}-tyramine and \textit{m}-hpaa in the human body remain unclear. Both of these metabolites can cross the blood-brain barrier (38, 39). Whereas the microbial metabolite \textit{p}-tyramine, which can be derived from tyrosine decarboxylation, has a number of biological effects and known targets (40, 41), \textit{m}-tyramine has been less frequently investigated. This metabolite inhibits dopamine synthesis in the rat striatum and is thought to be responsible for the striking behavioral effects of \textit{m}-tyrosine in rats (42, 43). Less is known about potential effects of \textit{m}-hpaa, even though this metabolite has recently been correlated with human disease. Urinary \textit{m}-hpaa is one of a group of five metabolites that distinguishes depressed patients from healthy controls (44), and this metabolite significantly elevated in children with autism spectrum disorders (45). However, the relevance of previous work to the role of \textit{m}-tyramine and \textit{m}-hpaa in disease development and treatment is unclear. Understanding how these metabolites function in Parkinson’s disease treatment and beyond will require a molecular understanding of microbial dopamine dehydroxylation.

The relevance of gut microbial dopamine dehydroxylation could extend beyond L-dopa metabolism and Parkinson’s disease (Figure 3.1). Humans secrete \textit{m}-tyramine in urine under homeostatic conditions (46, 47). Urinary \textit{m}-tyramine levels are insensitive dietary changes and do not vary significantly over time in humans, strongly suggesting that urinary \textit{m}-tyramine excreted
in the absence of ingested L-dopa is of endogenous origin (47, 48). These data – along with the aforementioned observations that high levels of dopamine are produced in the gut, that catecholamines are present in the intestinal lumen, and that intestinal pathogens utilize dopamine for growth – raise the possibility that the dopamine dehydroxylation that was first discovered as part of microbial L-dopa metabolism could also involve endogenously produced dopamine. Thus, understanding the molecular basis of the elusive dopamine dehydroxylation would enhance our understanding of L-dopa metabolism while also providing insight into the increasingly appreciated connection between the microbiota and host nervous system.
A) Dopamine is one of three major catecholamine neurotransmitters produced in the human body.

B) Metabolism and physiological effects of dopamine. The Parkinson’s drug L-dopa can be decarboxylated by host or gut microbial processes into dopamine, a neurotransmitter with diverse physiological effects. Dopamine can also be endogenously derived. The microbial dehydroxylation of dopamine into m-tyramine is the major focus of the work described in this chapter. AADC stands for aromatic amino acid decarboxylase. TyrDC stands for tyrosine decarboxylase, an enzyme that was described in Chapter 1 of this thesis. UGT stands for UDP-glucuronosyltransferase.
Understanding the biological roles of dopamine dehydroxylation will require a greater understanding of the specific gut microbial strains and enzymes involved. Previous work has shed light on environmental strains and enzymes capable of reductive aromatic dehydroxylation, the type of chemistry involved in the para-dehydroxylation of dopamine. At the time we began our work, 4-hydroxybenzoyl-CoA reductase (4-HBCR) from *Thauera aromatica* was the only biochemically characterized aromatic dehydroxylase. This molybdenum-dependent enzyme belongs to the xanthine oxidase family and reductively removes the para-hydroxyl group from 4-hydroxybenzoyl CoA, a key intermediate in the anaerobic degradation of aromatic compounds (Figure 3.2) (49). The enzyme employs a Birch-like radical mechanism involving two sequential single-electron reductions to accomplish the energetically demanding aromatic dehydroxylation (50). The proposed radical intermediates are stabilized by the electron-withdrawing CoA group. 4-HBCR was suggested to catalyze dehydroxylation of catechol 3,4-dihydroxybenzoyl CoA in *T. aromatica*, but this suggestion was never evaluated with purified enzymes (51). It is worth noting that the dehydroxylation of dopamine may involve a reaction mechanism that differs from that of 4-HBCR due to the lack of an electron-withdrawing group and the presence of an additional electron-donating hydroxyl group. This possibility will be discussed in more detail in Chapter 5. At the time we began our work, no microbial isolate or enzyme from any habitat was known to dehydroxylate dopamine.

**Figure 3.2 Reaction performed by 4-hydroxybenzoyl CoA reductase.**
Prior to the beginning of the work, the *T. aromatica* 4-hydroxybenzoyl CoA reductase (4-HBCR) was the only enzyme demonstrated to catalyze reductive aromatic dehydroxylation in vitro.
Figure 3.2 (Continued).
This enzyme, which proceeds by a Birch-type radical mechanism, is a central player in the degradation or aromatic compounds in the environment. The generated benzoyl CoA can be further metabolized to provide carbon for growth.

3.2 Results

3.2.1 Screen for dopamine dehydroxylation among human gut microbes

To identify dopamine dehydroxylation bacteria, we first screened a collection of representative human gut microbes for their ability to perform this reaction in anaerobic cultures. We selected an initial set of 18 gut strains that were available in our laboratory and included a subset of the major bacterial phylogenetic groups in the gut, including Firmicutes, Bacteroidetes, and Actinobacteria. To rapidly screen for metabolism, we applied a previously reported colorimetric assay that detects catechols (52). This method was developed in 1937 to distinguish L-dopa from tyrosine and involves acidification of culture supernatants followed by addition of a solution of sodium nitrite and sodium molybdate, producing a distinct yellow-colored solution. Addition of sodium hydroxide basifies the solution and forms a pink complex whose absorbance maximum can be measured spectrophotometrically at 500 nm. The mechanism underlying color formation has not been elucidated, but it likely involves addition of nitrous acid to the catechol ring with molybdate stabilizing the colored complex (52). This method detects dopamine in a concentration-dependent manner, while m-tyramine gave no signal above the background in growth medium (Figure 3.3). However, none of the strains screened (Enterococcus faecalis MMH594, Enterococcus faecalis TX0104, Enterococcus faecium TX01330, Clostridium aspargiforme DSM15981, Flavonifractor plautii ATCC29863, Clostridium sp. ATCC BAA-442, Bifidobacterium longum spp. Infantis ATCC 15697, Flavonifractor plautii ATCC 49531, Clostridium ramosum DSM 1402, Akkermansia muciniphila ATCC BAA-835, Clostridium
* bartlettii ATCC BAA-827, Enterococcus raffinosus ATCC 49464, Eubacterium limosum ATCC 10825, Eggerthella lenta DSM2243, Faecalibacterium prausnitzii ATCC 27766, Eubacterium rectale ATCC 33656, Lactobacillus reuteri ATCC 23272, Parabacteroides distasonis ATCC 8503) displayed clear dopamine dehydroxylation, including in the L-dopa decarboxylating organism *E. faecalis*. This suggested that other, unknown gut bacteria likely dehydroxylate dopamine in the human gut microbiota.

**Figure 3.3** The Arnow method for colorimetric detection of dopamine in bacterial cultures. A) Schematic for colorimetric assay from harvesting of bacterial cultures to data collection. B) BHI medium with 500 µM dopamine (1) or without dopamine (2) subjected to the colorimetric assay. Dopamine produces a striking pink color that can be quantified spectrophotometrically as absorbance at 500 nm. C) Absorbance at 500 nm across a range of dopamine and *m*-tyramine concentrations in BHI medium subjected to the colorimetric assay. While there is a linear dose-dependent increase in absorbance with dopamine, there is no increase in the presence of *m*-tyramine. Data are from single replicates.
3.2.2 Enrichment culturing from a human fecal sample identifies the dopamine dehydroxylating organism *Eggerthella lenta* A2

Having failed to uncover any dopamine dehydroxylating strains in our initial screen, we instead turned to enrichment culturing, a strategy that has been successfully used to isolate xenobiotic-metabolizing gut bacteria (53). In this method, complex microbial communities are grown under conditions where a specific phenotype, such as the presence of a metabolic pathway, provides a significant growth advantage (54). Successive passaging of cultures (e.g. inoculation into fresh media) selects for metabolizing strains over non-metabolizing microbes who cannot take advantage of the provided nutrients (for example electron acceptors, carbon source, or nitrogen sources) (54). We recognized that catechol dehydroxylation, which is a reductive transformation, resembles the reductive dehalogenation of chlorinated aromatics, a reaction that enables anaerobic respiration in certain environmental bacteria (55). This led us to hypothesize that dehydroxylation of dopamine might also facilitate anaerobic respiration and promote growth of human gut bacteria. The potential connection between this activity and bacterial growth provided an opportunity for isolation of dopamine metabolizers.

Thus, we inoculated a stool sample from a human donor into a minimal growth medium containing 0.5 mM dopamine as the sole electron acceptor (Figure 3.4). Passaging over multiple generations enriched for active strains as assessed by the colorimetric assay for catechol dehydroxylation. Plating the active cultures on rich MEGA medium eventually identified a strain of the gut Actinobacterium *Eggerthella lenta* (referred to herein as strain A2) that selectively removes the *para* hydroxyl group of dopamine to give *m*-tyramine (Figures 3.4 and 3.5). Interestingly, we had evaluated a different strain of *E. lenta* (DSM2243) for dopamine dehydroxylation in our original screen described in Section 3.2.1. The failure of that strain to
metabolize dopamine suggested potential strain-level variability in this activity. Additionally, certain *E. lenta* strains also inactivate the cardiac drug digoxin, suggesting a wider role for this gut organism in drug metabolism than previously appreciated (56, 57).

![Figure 3.4](image)

**Figure 3.4 Enrichment culturing for isolating dopamine dehydroxylating strains from complex human stool samples.**

A) General overview of enrichment culturing strategy. A human stool sample was inoculated into an anaerobic minimal medium containing 500 µM dopamine as the electron acceptor and 10 mM acetate as the electron donor. This sample was initially incubated at 37 °C for five days and was then passaged twice by a 1:100 dilution into fresh medium. Each successive passage was incubated for 48 hours at 37 °C. The last passage (named “enrichment endpoint” in panel B)) was then plated, allowing for isolation of active strains (red in the schematic). B) Pie charts describing the bacterial genus abundance at various points of enrichment culturing as assessed by 16S rRNA sequencing (>1% abundance). Following the plating of the culture at the enrichment endpoint, we picked what appeared to be individual colonies to obtain pure cultures. However, the initially picked colony contained two organisms: *Enterococcus* sp. (black) and *Eggerthella* sp. (purple), which were present at the end of enrichment culturing. We likely obtained two strains because the colonies grown on minimal media plates were extremely small and difficult to distinguish. Thus, we plated this co-culture on rich MEGA medium to better distinguish individual colonies, allowing us to obtain a pure culture of the metabolizing organism *Eggerthella lenta*. The isolated, active dopamine dehydroxylating strain was named *Eggerthella lenta* A2.
Figure 3.5 *Eggerthella lenta* A2 quantitatively and regiospecifically dehydroxylates dopamine to *m*-tyramine in anaerobic culture.

*E. lenta* A2 was grown with or without 500 µM dopamine for 48 hours anaerobically in BHI medium at 37 °C, and the culture supernatants were analyzed by LC-MS. Data represent LC-MS/MS ion chromatograms for simultaneous detection of dopamine and *m*-tyramine after 48 hours of anaerobic growth. Peak heights show the relative intensity of each mass and all chromatograms are shown on the same scale. Dopamine was completely dehydroxylated to give *m*-tyramine but not its regioisomer *p*-tyramine.
3.2.3 Dopamine dehydroxylation is inducible in *Eggerthella lenta* A2

To identify the dopamine dehydroxylating enzyme, we first searched the *E. lenta* A2 genome for genes encoding homologs of the only characterized aromatic *para*-dehydroxylase, the *T. aromatica* 4-HBCR (58). However, we found no hits, suggesting that this chemically challenging catechol dehydroxylation reaction likely involves a distinct enzyme. We did not have another clear enzyme candidate, preventing us from using genome searches alone to identify candidate enzymes, and prompting us to consider an alternative strategy. Previous work on the cardiac drug digoxin used RNA-sequencing to identify genes encoding drug-metabolizing enzymes in *E. lenta*. In these experiments, Haiser and co-workers revealed that *E. lenta* upregulated a two-gene operon more than 100-fold in response to the cardiac drug digoxin compared to a vehicle control (56). One of these genes was subsequently heterologously expressed, and in vitro assays confirmed that this enzyme (named the cardiac glycoside reductase 1, or Cgr1) specifically metabolized digoxin (57). Because *E. lenta* harbored many reductases that could possibly perform the conjugate reduction of digoxin, RNA-seq was critical to narrow down the list of candidate enzymes. We decided to apply this approach to identify the dopamine dehydroxylase enzyme in *E. lenta* A2.

First, we investigated whether dopamine dehydroxylation was an inducible activity in this organism. We grew the *E. lenta* A2 strain in the presence and absence of dopamine and then re-suspended each culture during mid-exponential phase in fresh medium with and without dopamine, allowing for further growth. This experiment revealed that cultures previously grown with dopamine completely consumed dopamine within two hours of re-suspension, whereas those previously grown without dopamine barely consumed this substrate (<20%) after three hours (Figure 3.6). While these data suggested dopamine dehydroxylation was inducible, they did not
distinguish whether the inducibility reflected differential transport into the cell or differential enzymatic activity. To further understand the inducibility, we assayed the dopamine dehydroxylase activity of cell lysates from cultures grown with and without dopamine. These assays suggested that the enzymatic activity was induced by dopamine and confirmed that dopamine dehydroxylation required anaerobic condition (Figure 3.7). Taken together, these results suggested RNA-sequencing would be a viable approach for enzyme discovery.

![Graph showing dopamine remaining over time](image)

**Figure 3.6** Culture experiments establishes that dopamine dehydroxylation is inducible in *Eggerthella lenta* A2.

*E. lenta* A2 was growth with or without 500 µM dopamine anaerobically in BHI medium at 37 °C until mid-exponential phase, at which point the cultures were resuspended in fresh BHI medium with or without dopamine. Culture supernatants were analyzed for dopamine using the colorimetric assay for catechol detection. Data represent the mean ± the SEM of three biological replicates.
Fig. 3.7 Assays with *E. lenta* A2 lysates suggest dopamine dehydroxylation is an inducible and oxygen sensitive enzymatic activity.

*E. lenta* A2 was grown anaerobically either without or with 500 µM dopamine in BHI medium containing 1% (w/v) arginine and 10 mM formate. Cultures were spun down at OD$_{600}$=0.500, and cell pellets were washed twice with anaerobic, sterile PBS to remove dopamine and any *m*-tyramine that had accumulated in the culture. Cells were lysed in 50 mM Tris pH 8 followed by addition of 250 µM dopamine to the lysate. Lysates were incubated at room temperature for 12 hours and were analyzed for *m*-tyramine production by LC-MS/MS. Lysate assays were performed anaerobically unless otherwise indicated. Data represent the mean ± the SEM of three biological replicates.
3.2.4 Optimization of culture conditions for RNA-sequencing in *Eggerthella lenta* A2

To inform the RNA-sequencing experiment, we characterized dopamine dehydroxylation in greater detail. Arginine is the main source of carbon, nitrogen, and energy in *E. lenta*, and previous work on digoxin metabolism suggests this nutrient can inhibit xenobiotic metabolism by suppressing transcription of metabolizing enzymes (56, 57, 59). We found that arginine increased *E. lenta* growth in a concentration-dependent manner but only inhibited dopamine dehydroxylation at concentrations above 1% (w/v) (Figure 3.8). This concentration was higher than previously observed with digoxin, which completely inhibited digoxin reduction at 0.5% arginine (w/v) (56). Additionally, unlike earlier work with digoxin, we only observed 41% inhibition of metabolism compared to the complete inhibition observed previously (Figure 3.8) (56).

In our initial experiments profiling the time course of metabolism, we added dopamine at the time of inoculation. However, in previous RNA-seq experiments in *E. lenta*, compound was spiked in during mid-exponential phase rather than at the time of inoculation (56, 60). When we added dopamine to exponentially growing cultures, we found that 1% arginine completely inhibited dehydroxylation (Figure 3.8). This finding presented a problem because without added arginine there might not be sufficient cell mass for successful RNA extraction and sequencing. However, the addition of 10 mM sodium formate, a molecule known to enhance metabolism of the dietary polyphenol catechin in *E. lenta* (61), to our cultures overcame the inhibitory effect of arginine and allowed us to observe dehydroxylation of dopamine when added to exponential cultures. In addition to its effect on dopamine dehydroxylation, formate significantly accelerated growth of *E. lenta*, a normally fastidious organism (Figure 3.7). Taken together, our results show that formate promotes growth without negatively affecting dopamine dehydroxylation in *E. lenta* and that the effects of arginine on this metabolism depend on when dopamine is added to the
medium. These results also informed final conditions for RNA-sequencing, in which we grew *E. lenta* with 1% arginine and 10 mM formate, added dopamine at early exponential phase, and harvested the cells in mid-exponential phase. The mechanism underlying the growth-promoting effect of formate is unknown, but one possible explanation is direct metabolism of formate. *E. lenta* harbors genes predicted to encode for formate dehydrogenase, a membrane-associated enzyme that accelerates growth by coupling formate oxidation to proton translocation (62-65). Alternatively, formate could be incorporated into the Wood-Ljungdahl pathway (66, 67). This pathway forms the cornerstone of acetogenesis, a type of microbial metabolism that reduces carbon dioxide and oxidizes hydrogen for carbon and energy harvest (66) *E. lenta* genetically encodes for key steps of this pathway and is predicted to be an acetogen (62).

Figure 3.8 Impact of arginine on growth and dopamine dehydroxylation in *E. lenta* A2 when dopamine is added at the time of inoculation. *E. lenta* A2 was grown anaerobically either without or with 500 µM dopamine in BHI medium with increasing concentrations of arginine (w/v) for 48 hours. A) Endpoint growth was measured spectrophotometrically and B) dopamine dehydroxylation was assessed using a colorimetric assay for catechol detection. Data represent the mean ± the SEM of three biological replicates.
Figure 3.9 Impact of arginine and formate dopamine dehydroxylation E. lenta A2 when dopamine is added during exponential growth.

E. lenta A2 was grown anaerobically either BHI medium with 1% arginine, 1% arginine and 10 mM formate, or no additive. Dopamine was added at the time indicated by the arrow and metabolism was measured using LC-MS/MS. A) Arginine boosts growth of E. lenta relative to no additive, and the addition of formate to arginine-containing cultures increases the growth rate and maximum growth further. B) Dopamine added to exponentially growing cultures in BHI medium containing 1% arginine and 10 mM formate is rapidly dehydroxylated into m-tyramine. C) Dopamine added to exponentially growing cultures in BHI medium containing 1% arginine is not dehydroxylated into m-tyramine, indicating that arginine inhibits dehydroxylation in the absence of formate. Data represent the mean ± the SEM of three biological replicates.
3.2.5 RNA-sequencing identifies a putative molybdenum-dependent dopamine dehydroxylase

Having optimized the culture conditions, we performed an RNA-sequencing experiment to identify the dopamine dehydroxylating enzyme in *E. lenta* A2. We grew *E. lenta* A2 in BHI medium containing 1% arginine and 10 mM formate. We then added dopamine in early exponential phase (OD$_{600}$ = 0.200), and harvested cells when approximately 30% of the substrate had been dehydroxylated as assessed by the colorimetric assay. A set of cultures where vehicle (water) was added served as the control. This experiment revealed upregulation of two distinct genomic loci in the strains grown with dopamine, including >2,500-fold upregulation of 3 co-localized genes (Figure 3.10 and Table 3.1). These highly upregulated genes encode a predicted bis-molybdopterin guanine dinucleotide cofactor (bis-MGD)-containing enzyme belonging to the DMSO reductase family, with one of the genes predicted to encode the Bis-MGD-binding catalytic subunit and the other two predicted to encode a membrane anchor and electron shuttling ferredoxin (68). Importantly, only one bis-MGD enzyme-encoding gene was upregulated in response to dopamine (Table 3.1).

Bis-MGD enzymes catalyze a wide variety of oxygen-transfer reactions, including DMSO reduction, acetylene hydration, and nitrate reduction, leading us to hypothesize that this enzyme was the dopamine dehydroxylase (Dadh) (69). The biosynthesis of the bis-MGD cofactor, as well as the chemistry and maturation of bis-MGD enzymes, is discussed in detail in Chapter 5 of this thesis. The phylogenetic diversity of the bis-MGD enzyme family is discussed in Chapter 4. As *E. lenta* is genetically intractable, we could not employ classical genetic knockout experiments to verify this functional assignment. Instead, we turned to chemical genetics, exploring whether the dopamine dehydroxylation activity was molybdenum-dependent by culturing *E. lenta* A2 in the
presence of tungstate. Substitution of molybdate with tungstate during bis-MGD biosynthesis generates an inactive metalloc cofactor (Figure 3.11) (70). Thus, tungstate and molybdate have been used as tools in microbiology and biochemistry to probe the metal and redox dependence of specific metabolic pathways and to inhibit molybdenum-dependent gut microbial metabolism (70-72). For example, tungstate addition inhibits the gut bacterial molybdenum-dependent reduction of nitrate in pure cultures and in mice (70). We hypothesized that tungstate similarly would inhibit dopamine dehydroxylation by *E. lenta* if a bis-MGD enzyme was involved. Consistent with this prediction, treating cultures of *E. lenta* A2 with tungstate inhibited dopamine dehydroxylation without affecting growth (Figure 3.12). In contrast, molybdate did not inhibit dopamine dehydroxylation. Incubating cell lysates with tungstate had no effect, which is consistent with inhibition by tungstate requiring active bis-MGD biosynthesis (Figure 3.13) (73). Taken together, these data implicate the molybdenum-dependent Dadh in dopamine dehydroxylation by *E. lenta* A2.
Table 3.1 Genes upregulated in *E. lenta* A2 in response to dopamine. Data represent genes upregulated in *E. lenta* A2 when grown with 500 µM dopamine relative to a vehicle control in BHI medium containing 1% (w/v) arginine and 10 mM formate. The three subunits predicted to be involved in dopamine dehydroxylase are highlighted in red.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>237-Elenta-A2_00534</td>
<td>4Fe-4S dicluster domain-containing protein (electron shuttling ferredoxin subunit)</td>
<td>11.43265558</td>
<td>1.11E-26</td>
</tr>
<tr>
<td>237-Elenta-A2_00535</td>
<td>Molybdopterin dinucleotide-binding protein (catalytic subunit)</td>
<td>11.32638074</td>
<td>7.46E-79</td>
</tr>
<tr>
<td>237-Elenta-A2_00533</td>
<td>Hypothetical protein (membrane anchor subunit)</td>
<td>10.7757051</td>
<td>1.60E-47</td>
</tr>
<tr>
<td>237-Elenta-A2_00540</td>
<td>Uncharacterized component of anaerobic dehydrogenase</td>
<td>9.648691059</td>
<td>7.67E-19</td>
</tr>
<tr>
<td>237-Elenta-A2_00539</td>
<td>Hypothetical protein</td>
<td>9.064741704</td>
<td>1.09E-16</td>
</tr>
<tr>
<td>237-Elenta-A2_00542</td>
<td>4Fe-4S binding protein</td>
<td>7.74321557</td>
<td>4.18E-18</td>
</tr>
<tr>
<td>237-Elenta-A2_00543</td>
<td>4Fe-4S ferredoxin</td>
<td>7.452293617</td>
<td>2.49E-44</td>
</tr>
<tr>
<td>237-Elenta-A2_00541</td>
<td>4Fe-4S ferredoxin</td>
<td>7.108180453</td>
<td>6.16E-20</td>
</tr>
<tr>
<td>237-Elenta-A2_00538</td>
<td>Hypothetical protein</td>
<td>5.687034137</td>
<td>4.53E-109</td>
</tr>
<tr>
<td>237-Elenta-A2_00544</td>
<td>Hypothetical protein</td>
<td>4.974609256</td>
<td>5.23E-48</td>
</tr>
<tr>
<td>237-Elenta-A2_02490</td>
<td>Energy-coupling factor ABC transporter ATP-binding protein</td>
<td>2.430668122</td>
<td>5.71E-28</td>
</tr>
<tr>
<td>237-Elenta-A2_02489</td>
<td>ABC transporter ATP-binding protein</td>
<td>2.079536895</td>
<td>2.58E-23</td>
</tr>
<tr>
<td>237-Elenta-A2_02488</td>
<td>Energy-coupling factor transporter transmembrane protein EcT</td>
<td>1.749835331</td>
<td>1.58E-18</td>
</tr>
<tr>
<td>237-Elenta-A2_02491</td>
<td>Hypothetical protein</td>
<td>1.601691749</td>
<td>2.21E-20</td>
</tr>
<tr>
<td>237-Elenta-A2_01052</td>
<td>TrpB-like pyridoxal phosphate-dependent enzyme</td>
<td>1.228094756</td>
<td>7.69E-52</td>
</tr>
</tbody>
</table>
Figure 3.10 RNA-Seq identifies a putative bis-MGD-dependent dopamine dehydroxylase (Dadh) in *E. lenta* A2.

Differentially expressed candidate genes (FDR<0.1 and fold-change (FC)>2) are plotted as a function of genome position revealing three discrete loci of differentially expressed genes. (Inset) analysis of the largest cluster of differentially expressed genes at 0.665 megabase pairs (Mbp) in the scaffolded assembly (190 Kbp in the reference contig) revealed the putative *dadh* up-regulated by 2,568-fold in response to dopamine. The figure was prepared by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
Figure 3.11 Model for tungstate inhibition of dopamine dehydroxylation in *E. lenta* A2.

A) Tungstate (WO$_4^{2-}$) or molybdate (MoO$_4^{2-}$) can be incorporated into the molybdenum dinucleotide cofactor during its complex biosynthesis, producing a cofactor where M is either W (tungsten) or Mo (molybdenum). This cofactor is then installed intact into its cognate bis-MGD-dependent enzyme, such as the dopamine dehydroxylase (Dadh). B) We hypothesized that the incorporation of tungsten into the cofactor would prevent dopamine dehydroxylation by Dadh, while molybdenum should not impact dehydroxylation. MPT = molybdopterin.
**Figure 3.12** Tungstate inhibits dopamine dehydroxylation without inhibiting growth of *E. lenta* A2.

A) Tungstate treatment inhibits dehydroxylation of dopamine by *E. lenta* A2. Cultures were grown anaerobically with tungstate (WO$_4^{2-}$) or molybdate (MoO$_4^{2-}$) for 48 hours with 0.5 mM dopamine. Bar graphs represent the mean ± the SEM of three biological replicates.

B) *E. lenta* A2 was grown anaerobically in BHI medium at 37 °C with varying concentrations of tungstate. There was no obvious dose-dependent growth defect in the presence of tungstate. The data shown are the mean of three replicate growth experiments ± the SEM.
Figure 3.13 Tungstate does not inhibit dopamine dehydroxylation in cell lysates.

*E. lenta* A2 was grown either without or with 500 µM dopamine (DA) in BHI medium containing 1% (w/v) arginine and 10 mM formate. Cultures were spun down at OD$_{600}$ = 0.500 and cell pellets were washed twice with anaerobic, sterile PBS to remove dopamine and any m-tyramine that had accumulated in the culture. Pellets were lysed anaerobically in 50 mM Tris pH 8 followed by addition of 500 µM dopamine to lysates. Lysates were left at room temperature for 12 hours and were analyzed for m-tyramine production by LC-MS/MS. Peak heights show the relative intensity of each mass and all chromatograms are shown on the same scale. There was no dose-dependent effect of tungstate on dopamine dehydroxylation, while the addition of oxygen (O$_2$) completely blocked this activity. Dopamine = DA.
3.2.6 Native purification of the dopamine dehydroxylase from *E. lenta* A2

To biochemically confirm the identity of the dopamine-dehydroxylating enzyme from *E. lenta* A2 we first attempted heterologous expression. However, as will be detailed in Chapter 5, expression of multiple constructs in multiple phylogenetically diverse hosts failed to yield active, soluble protein. Instead, we turned to native purification, a strategy commonly used to access molybdenum-dependent enzymes like Dadh \(^{74-76}\). We began by establishing the location of activity in anaerobically lysed cells, using LC-MS to track activity. The genomic context of Dadh suggested that Dadh may be associated with a membrane anchor like other molybdenum-dependent enzymes (Figure 3.14) \(^{68}\), leading us to test whether this activity was membrane-associated in *E. lenta* A2. However, following cell lysis via sonication and ultracentrifugation of the soluble fraction to pellet membranes, we found that the majority of activity remained in the supernatant under these conditions (Figure 3.15). It is possible that the protein is membrane-associated in vivo but that the sonication step dissociates the complex.

Given this observation, we pursued purification of the dopamine dehydroxylase from the soluble fraction of cell lysates. As an initial clean-up step of the soluble fraction, we used ammonium sulfate precipitation. A majority of the dopamine dehydroxylase activity precipitated at 40% ammonium sulfate (Figure 3.16). Following optimization of Fast Protein Liquid Chromatography (FPLC) conditions, we developed a protocol in which the resuspended ammonium sulfate pellet was subjected to activity-based anaerobic fractionation on hydrophobic interaction chromatography, anion exchange chromatography, and finally size exclusion chromatography. Applying this purification procedure to 4 L of *E. lenta* A2 culture yielded a dopamine dehydroxylating fraction containing four major proteins as assessed by SDS-PAGE (Figures 3.17 and 3.18 and table 3.2). Dehydroxylation activity at the different stages of
purification correlated with the intensity of a 115 kDa band. This band was confirmed to be the catalytic subunit of Dadh using mass spectrometry (Figure 3.17). Importantly, Dadh was the only isolated protein observed to be upregulated in the presence of dopamine in the previous RNA-Seq experiment (Tables 3.1 and 3.2). Together, these data strongly support the assignment of this enzyme as the dopamine dehydroxylase.

Figure 3.14 Predicted model for *E. lenta* A2 Dadh.

The predicted catalytic molybdenum-containing subunit (a, red) is co-localized with a predicted 4Fe-4S cluster-containing protein (b, purple) and a predicted membrane anchor (c, gold) in the genome. The molybdenum-containing subunit is also predicted to contain an N-terminal 4Fe-4S cluster. The catalytic subunit carries a Twin-Arginine-Translocation (TAT) signal sequence, suggesting it is exported. There, together with the predicted 4Fe-4S cluster-containing protein it could form a membrane-anchored complex that receives electrons from quinone oxidation in the membrane. Electrons released from quinone oxidation could move through the iron sulfur clusters of the two subunits to ultimately reduce the molybdenum cofactor that performs the 2e- reduction of the catechol substrate.
**Figure 3.15 Dopamine dehydroxylase activity is not membrane-associated in *E. lenta* lysates.** Soluble fractions of *E. lenta* A2 lysates grown with dopamine were ultracentrifuged to pellet membranes. The cells had been lysed in 50 mM Tris pH 8 containing 0.1% Triton as a detergent. The membrane pellet and supernatants were subjected to an activity assay in which 0.5 mM dopamine, 1 mM methyl viologen, and 2 mM sodium dithionite were added. Reactions were incubated anaerobically at room temperature for 14 hours and metabolites were analyzed by LC-MS/MS. This experiment revealed that the activity was not membrane-associated in *E. lenta* A2 following sonication to lyse cells.
Figure 3.16 Ammonium sulfate precipitation of dopamine dehydroxylase activity.
Soluble fractions of *E. lenta* A2 lysates grown with dopamine were subjected to varying concentrations of ammonium sulfate for 1.5 hours. The cells had been lysed in 50 mM Tris pH 8. After 1.5 hours, the mixtures were spun down. The resuspended pellet and the supernatant at each concentration of ammonium sulfate were subjected to an activity assay in which 0.5 mM dopamine, 1 mM methyl viologen, and 2 mM sodium dithionite were added. Reactions were incubated anaerobically at room temperature for 14 hours and metabolites were analyzed by LC-MS/MS. A) results from the pellets B) results from the supernatants. This experiment indicated that the activity precipitated at 40% ammonium sulfate (w/v) onwards.
Figure 3.17 Native purification of the dopamine dehydroxylase from *E. lenta* A2.
A) Schematic of purification procedure. The active fractions at each step were subjected to further chromatography.
Figure 3.17 (Continued).
The active size exclusion chromatography fractions were subjected to SDS-PAGE analysis and proteomics. B) SDS-page of size exclusion chromatography fractions from activity-based purification of the dopamine dehydroxylase from *E. lenta* A2. Ladder is the Precision Plus Protein™ All Blue Standards (first lane from the left), while the subsequent lanes represent fractions from the size exclusion column, the last chromatography step of the activity-based purification from *E. lenta* A2. Each fraction was incubated for 12-14 hours anaerobically with 500 µM dopamine, 1 mM sodium dithionite, and 2 mM each of the electron donors benzyl viologen, methyl viologen, and diquat dibromide in 50 mM Tris pH 8 buffer at room temperature. After incubation, enzyme assay mixtures were analyzed by LC-MS/MS for dopamine and *m*-tyramine. The bar graphs represent the total dehydroxylation by each fraction (lanes 1-5), and this value was calculated as the concentration of *m*-tyramine normalized by the total concentration of *m*-tyramine and dopamine. The red asterisk indicates the band corresponding to Dadh as confirmed by proteomics. This band tracks with activity, unlike other protein contaminants. The fraction shown in lane 5 was used for global proteomics, and the gel band representing the dopamine dehydroxylase (red asterisk) across lanes 4 and 5 was also cut out and subjected to proteomics to confirm this band’s identity individually.
Figure 3.18 Dadh-containing fractions purified from *E. lenta* A2 regiospecifically dehydroxylate dopamine to produce m-tyramine.

Extracted LC-MS/MS ion chromatograms for simultaneous detection of dopamine and m-tyramine after 12 hours of anaerobic incubation of enzyme preparation with 500 µM dopamine and artificial electron donors at room temperature. Peak heights show the relative intensity of each mass and all chromatograms are shown on the same scale.

Table 3.2 Proteomics identification of Dadh.
Proteomics identification of the major bands in active size exclusion fraction (lane 5 in Figure 3.15) during purification of the dopamine dehydroxylase from *E. lenta* A2. The catalytic subunit of the dopamine dehydroxylase is highlighted in red.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>Coverage</th>
<th># AAs</th>
<th>MW [kDa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>237-Elenta-A2_00535</td>
<td>Molybdopterin dinucleotide binding region</td>
<td>58.41</td>
<td>1017</td>
<td>115</td>
</tr>
<tr>
<td>237-Elenta-A2_00242</td>
<td>Heat shock protein Hsp90</td>
<td>31.42</td>
<td>646</td>
<td>72.6</td>
</tr>
<tr>
<td>237-Elenta-A2_02071</td>
<td>Pyruvate ferredoxin/flavodoxin oxidoreductase</td>
<td>33.47</td>
<td>1180</td>
<td>127.6</td>
</tr>
<tr>
<td>237-Elenta-A2_01108</td>
<td>Nickel-dependent hydrogenase large subunit</td>
<td>5.96</td>
<td>554</td>
<td>60.7</td>
</tr>
</tbody>
</table>
3.2.7 A single nucleotide polymorphism (SNP) discriminates dopamine metabolizing *E. lenta* strains from non-metabolizers

We next assessed whether the presence of *dadh* in microbial genomes correlated with dopamine dehydroxylation. A BLASTP search with stringent parameters to capture closely related homologs (100% query cover, >80% amino acid identity) in the NCBI non-redundant protein sequence database revealed that this enzyme is restricted to *E. lenta* and its close Actinobacterial relatives (Table 3.3), prompting us to screen a collection of gut Actinobacterial isolates (60, 77) for their ability to dehydroxylate dopamine in anaerobic culture. This collection harbors 26 different gut bacterial members of the Coriobacteriia class, including strains of *Eggerthella*, *Gordonibacter*, and *Paraeggerthella*, that were isolated from human stool samples in North America, Europe, and Asia. Dadh appeared to be encoded by all strains except the two *Gordoinibacter* (24 of the 26 strains total; 92–100% amino acid ID; Figure 3.19 and Table 3.4). However, only 10 *Eggerthella* strains quantitatively converted dopamine to *m*-tyramine, with low (<11%) or no metabolism in the others (Figure 3.20). This strain-level variability in dopamine metabolism reinforces that gut microbial species identity is often not predictive of metabolic functions (49, 50). This finding is also consistent with previous studies of digoxin metabolism by this bacterial species. However, in contrast to digoxin metabolism, where enzyme presence perfectly correlated with metabolism across gut Actinobacteria, our data indicated that the putative metabolizing enzyme was present even in non-metabolizing strains (57).

To better understand this variation in metabolism, we performed RNA-sequencing experiments with metabolizing (*E. lenta* 28b) and non-metabolizing (*E. lenta* DSM2243) strains in the presence and absence of dopamine. Surprisingly, *dadh* was upregulated in response to dopamine in both strains, indicating that lack of activity in *E. lenta* DSM2243 did not arise from
differences in transcription (Tables 3.5 and 3.6). Aligning the Dadh protein sequences, we instead
found a single amino acid substitution that correlated with metabolizer status: position 506 is an
arginine in metabolizing strains and a serine in inactive strains (Figure 3.20 and Figure 3.21). This
change arises from a single nucleotide polymorphism (SNP) in dadh, with the codon being CGC
for arginine and AGC for serine. The only exception, *E. lenta* W1BHI6, which is an inactive strain,
has the Arg506 variant and an additional substitution nearby (Cys500) (Figure 3.20). Thus, specific
amino acid residues in the Dadh enzyme, rather than presence or transcription of *dadh*, predict
dopamine dehydroxylation among gut bacterial strains. These Dadh variants did not correlate with
*E. lenta* phylogeny (Figure 3.20), suggesting that this activity has been gained/lost multiple times.
Table 3.3 Predicted dopamine dehydroxylases among sequenced isolates.
List of predicted dopamine dehydroxylases as retrieved from the non-redundant NCBI database, which only includes a subset of members of the members of our Actinobacterial strain library, at a threshold of 80% amino acid identity.

<table>
<thead>
<tr>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Annotation [species]</th>
<th>Query cover</th>
<th>E-value</th>
<th>% ID</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Eggerthella lenta]</td>
<td>100%</td>
<td>0</td>
<td>100%</td>
<td>WP_086414988.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Eggerthella lenta DSM2243]</td>
<td>100%</td>
<td>0</td>
<td>99%</td>
<td>WP_015759999.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Eggerthella lenta DSM2243]</td>
<td>100%</td>
<td>0</td>
<td>99%</td>
<td>WP_009608130.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Eggerthella lenta DSM2243]</td>
<td>100%</td>
<td>0</td>
<td>99%</td>
<td>WP_009306100.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Eggerthella lenta DSM2243]</td>
<td>100%</td>
<td>0</td>
<td>98%</td>
<td>WP_101721531.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Gordonibacter sp. An232A]</td>
<td>100%</td>
<td>0</td>
<td>94%</td>
<td>WP_087197552.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Rubneribacter badeniensis]</td>
<td>100%</td>
<td>0</td>
<td>93%</td>
<td>WP_103263115.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Gordonibacter sp. An230]</td>
<td>100%</td>
<td>0</td>
<td>93%</td>
<td>WP_087194795.1</td>
</tr>
<tr>
<td>Coriobacteria</td>
<td>Eggerthellales</td>
<td>Eggerthellaceae</td>
<td>molybdopterin dinucleotide-binding protein [Slackia heliotrinireducens]</td>
<td>100%</td>
<td>0</td>
<td>82%</td>
<td>WP_012799829.1</td>
</tr>
</tbody>
</table>
Figure 3.19 Global alignment of the dopamine dehydroxylase locus across an Actinobacterial library.

Reads from genome sequencing were mapped to the reference *E. lenta* A2 genome contig containing *dadh* and surrounding sequences using Bowtie2 and filtered for a minimum mapping quality of 10. Variants were called when >80% of reads supported an alternate sequence. Indel = insertion or deletion. Subst = amino acid substitution relative to the *E. lenta* A2 reference. Log2(FC) = log2fold change in gene expression in response to 500 µM dopamine relative to a vehicle control in *E. lenta* A2. El = *Eggerthella lenta*, Es = *Eggerthella sinensis* and Ph = *Paraeggerthella hongonensis*. This analysis and the figure were prepared by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
Table 3.4 Dopamine dehydroxylases in gut Actinobacteria.
List of predicted dopamine dehydroxylases retrieved from a custom genome database of 28 Actinobacterial isolates at a threshold of 92% amino acid identity.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Query cover</th>
<th>E-value</th>
<th>% Amino Acid ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSM11767</td>
<td>100.00%</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>AB12n2</td>
<td>100.00%</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>A2</td>
<td>100.00%</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>RC46F</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>MRn12</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>DSM18163</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>DSM15644</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>CC86D54</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>AN51LG</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>AB8n2</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>28b</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>22C</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>16A</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>14A</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>11c</td>
<td>100.00%</td>
<td>0</td>
<td>99.9%</td>
</tr>
<tr>
<td>HGA1</td>
<td>100.00%</td>
<td>0</td>
<td>99.7%</td>
</tr>
<tr>
<td>1160FAA</td>
<td>100.00%</td>
<td>0</td>
<td>99.6%</td>
</tr>
<tr>
<td>1160AFAA</td>
<td>100.00%</td>
<td>0</td>
<td>99.6%</td>
</tr>
<tr>
<td>326IFAA</td>
<td>100.00%</td>
<td>0</td>
<td>99.6%</td>
</tr>
<tr>
<td>1356FAA</td>
<td>100.00%</td>
<td>0</td>
<td>98.9%</td>
</tr>
<tr>
<td>Valencia</td>
<td>100.00%</td>
<td>0</td>
<td>98.3%</td>
</tr>
<tr>
<td>CC82BHI2</td>
<td>100.00%</td>
<td>0</td>
<td>98.3%</td>
</tr>
<tr>
<td>CC75D52</td>
<td>100.00%</td>
<td>0</td>
<td>98.3%</td>
</tr>
<tr>
<td>W1BHI6</td>
<td>100.00%</td>
<td>0</td>
<td>98.1%</td>
</tr>
<tr>
<td>RC22A</td>
<td>100.00%</td>
<td>0</td>
<td>93.2%</td>
</tr>
<tr>
<td>DSM16107</td>
<td>100.00%</td>
<td>0</td>
<td>92.1%</td>
</tr>
</tbody>
</table>
A single amino acid variant predicts dopamine metabolism by *E. lenta* and related strains and does not correlate with phylogeny.

Strains were cultured anaerobically with 500 µM dopamine in BHI for 48 hours (*El* *E. lenta*, *Es* *E. sinensis*, *Gs* *Gordonibacter* sp., *Gp* *Gordonibacter pamelaeae*). High (100% conversion) and low (<11% conversion) metabolizers are denoted in red and blue. For each strain, data points represent biological replicates. * P<0.05 ANOVA with Dunnett’s test vs sterile controls.
Table 3.5 Genes upregulated in *E. lenta* DSM2243 when grown with 500 µM dopamine.
Data represent the genes upregulated in *E. lenta* DSM2243 when grown with 500 µM dopamine relative to a vehicle control in BHI medium containing 1% (w/v) arginine and 10 mM formate. The catalytic subunit of the dopamine dehydroxylase is highlighted in red.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2 Fold Change</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggerthella_lenta_DSM2243_00482</td>
<td>Molybdopterin dinucleotide binding region</td>
<td>3.641149985</td>
<td>1.44E-163</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00480</td>
<td>Hypothetical protein</td>
<td>3.140977549</td>
<td>3.76E-119</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00481</td>
<td>4Fe-4S ferredoxin</td>
<td>2.73148694</td>
<td>2.63E-89</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00485</td>
<td>Hypothetical protein</td>
<td>1.668195003</td>
<td>7.10E-34</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00488</td>
<td>4Fe-4S ferredoxin</td>
<td>1.478077832</td>
<td>2.12E-27</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00489</td>
<td>Hypothetical protein</td>
<td>1.395767311</td>
<td>6.78E-25</td>
</tr>
<tr>
<td>Eggerthella_lenta_DSM2243_00582</td>
<td>Ferrous iron transport protein</td>
<td>1.097665786</td>
<td>1.74E-14</td>
</tr>
</tbody>
</table>

Table 3.6 Genes upregulated in *E. lenta* 28b when grown in the presence of dopamine.
Data represent the genes upregulated in *E. lenta* 28b when grown with 500 µM dopamine relative to a vehicle control in BHI medium containing 1% (w/v) arginine and 10 mM formate (log2foldchange>2). The catalytic subunit of the dopamine dehydroxylase is highlighted in red.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2 Fold Change</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggerthella_lenta_28b_02709</td>
<td>Molybdopterin dinucleotide binding region</td>
<td>5.339268092</td>
<td>6.01E-86</td>
</tr>
<tr>
<td>Eggerthella_lenta_28b_02711</td>
<td>Hypothetical protein</td>
<td>4.930759063</td>
<td>7.41E-65</td>
</tr>
<tr>
<td>Eggerthella_lenta_28b_02710</td>
<td>4Fe-4S ferredoxin</td>
<td>3.020280491</td>
<td>2.27E-19</td>
</tr>
<tr>
<td>Eggerthella_lenta_28b_02706</td>
<td>Hypothetical protein</td>
<td>2.878320413</td>
<td>3.28E-18</td>
</tr>
<tr>
<td>Eggerthella_lenta_28b_02701</td>
<td>4Fe-4S ferredoxin</td>
<td>2.611526487</td>
<td>2.50E-14</td>
</tr>
<tr>
<td>Eggerthella_lenta_28b_02708</td>
<td>Transcriptional regulator, LuxR family</td>
<td>2.06813492</td>
<td>9.27E-09</td>
</tr>
</tbody>
</table>
Alignment of Dadh protein sequences from the Actinobacterial strain library.

Alignment of the dopamine dehydroxylase sequences was performed in Jalview version 2.10.4, allowing for identification of the amino acid residue at position 506. Positions 493 through 511 are shown. Red asterisks indicate the positions that may explain metabolizer status.

* Indicates position that may explain metabolizer status.
The dopamine dehydroxylase SNP status predicts dopamine dehydroxylation by complex gut microbiota communities

We next asked whether *E. lenta* and Dadh metabolized dopamine in complex gut microbiota communities. Incubating stool samples from 15 unrelated neurologically healthy humans ex vivo, we found that dopamine dehydroxylation separated into groups of high metabolizers and low metabolizers (Figure 3.22). As expected from our assays with isolated strains, qPCR assays on the gDNA extracted from these samples revealed that neither *dadh* nor *E. lenta* abundance explained the observed patterns of low and high metabolism (Figure 3.22). However, when we amplified *dadh* from these cultures and determined the SNP status at position 506, we found samples containing the Arg506 variant quantitatively metabolized dopamine, while the activity of samples carrying the Ser506 variant was indistinguishable from the non-metabolizing *E. lenta* DSM2243 strain (Figure 3.22). Samples with both SNPs present were not included in this analysis.

Further metagenomic analysis by Dr. Jordan Bisanz (Turnbaugh lab, UCSF), revealed that *dadh* was highly prevalent (>70%) in gut microbiomes from human subjects and the two *dadh* variants were present among this population (Figures 3.23 and 3.24). *Dadh* significantly correlated with *E. lenta* abundance, indicating that this is likely a major dehydroxylation organisme dopamine in the human gut. *Dadh* was also more prevalent than *tyrDC*, the gene responsible for L-dopa decarboxylation (Figure 3.23). Finally, we found that the *dadh* SNP predicted dopamine dehydroxylation in Parkinson’s patient stool sample cultures, underscoring the clinical relevance of our results (Figure 3.25). Taken together, these findings indicate a single amino acid residue in a gut microbial enzyme correlates with dopamine metabolism in complex communities. Given that both SNPs of this enzyme are widely distributed among human patients, we speculate that SNPs
may influence xenobiotic metabolism in the context of not only the host genome (78), but also in the human gut microbiome (79).

Figure 3.22 The Dadh SNP predicts dopamine dehydroxylation by complex gut microbiota samples from neurologically healthy patients ex vivo.
A) Dopamine dehydroxylation by gut microbiota samples of 15 unrelated individuals. Samples were cultured for 48 hours with 0.5 mM dopamine in BHI medium containing 1 % arginine and 10 mM formate. Bars are mean ± the SEM of n=6 for low reducers (<50%) and n=9 for high reducers (>50%). *** P=0.0002, one tailed Mann-Whitney test. B) Dadh abundance does not correlate with dehydroxylation by human gut microbiotas. Data represent qPCR with Dadh-specific primers. Each data point is the dadh abundance in each sample shown in A). Bars represent the mean and standard error. C) E. lenta abundance does not correlate with dehydroxylation by human gut microbiotas. Data represent qPCR with E. lenta 16S rRNA-specific primers. Each data point is the E. lenta abundance in each sample shown in A). Bars represent the mean and standard error. D) Dadh sequence variants predict dopamine dehydroxylation ex vivo. Full-length dadh from each culture in (A) was sequenced using primers specific for the region containing position 506. Samples where a mix of variants were present (n=3) were removed. Bars represent the mean and SEM (n=3 for AGC samples, n=7 for CGC samples, n=3 for DSM2243, and n=3 for A2). ** P=0.0083, one tailed Mann-Whitney test, CGC samples vs. AGC samples.
Figure 3.23 Metagenomic analysis of *E. lenta/dadh* and *E. faecalis/tyrDC* abundance and prevalence across human patients.

A) Correlation of gene/bacterial abundances clearly demonstrates strong linear correlations between *E. lenta/dadh* and *Enterococcus/tyrDC* (R\(^2\)>0.812, P<2.2e-16 Pearson’s Correlation). B) Prevalence estimates as a function of minimum abundance reveals that both *dadh* and *tyrDC* are highly prevalent in the human gut microbiome albeit at low relative abundances. The analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
Figure 3.24 SNP analysis of *dadh* in human gut metagenomes containing high-coverage *E. lenta* genomes.

The two amino acid variants of interest in Dadh: p.Ser500Cys p.Arg506Ser were found to be supported in metagenomes with the dominant SNP prevalence similar to those observed in the isolate strain collection (1/76 and 38/77 respectively). The analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
Figure 3.25 The Dadh SNP predicts dopamine dehydroxylation by complex gut microbiota samples Parkinson’s disease patients ex vivo.

A) LC-MS was used to quantify dopamine dehydroxylation in the fecal microbiotas of 12 unrelated Parkinson’s disease patients. Samples were cultured for 48 hours in BHI medium containing 1% (w/v) arginine, 10 mM formate, and 0.5 mM dopamine and metabolites were analyzed by LC-MS. A) Bars represent the mean and SEM of n=3 for low reducers (<50%) and n=9 for high reducers (>50%) ** P=0.0045, one tailed Mann-Whitney test, low reducers vs. high reducers. B) Dadh abundance does not correlate with dehydroxylation by fecal microbiomes from 12 unrelated Parkinson’s disease patients. Data represent qPCR with Dadh-specific primers. Each data point is the dadh abundance in each sample shown in A). Bars represent the mean and standard error. C) E. lenta abundance does not correlate with dehydroxylation by fecal microbiotas from 12 unrelated Parkinson’s patients. Data represent qPCR with E. lenta-specific primers. Each data point is the E. lenta abundance in each sample shown in A). Bars represent the mean and standard error. D) Dadh sequence variants predict dopamine dehydroxylation ex vivo. gDNA from each culture in (A) was extracted followed by PCR amplification of the dadh gene. The identities of the sequence variants were assessed by Sanger sequencing of dadh using primers specific for the region containing position 506. Samples where a mix of variants were present (n=6) were removed. Bars represent the mean and SEM (n=2 for AGC samples, n=4 for CGC samples). While all AGC samples were low metabolizers (<30%) and all CGC samples displayed quantitative metabolism (>98%), the difference in dehydroxylation between the groups did not reach statistical significance, likely due to the small number of samples involved (p=0.067, one-tailed Mann-Whitney test, n=2 AGC vs n=4 CGC).
3.2.9 *E. faecalis* and *E. lenta* produce *m*-tyramine from L-dopa in co-culture and in complex gut microbiota communities

Our studies of L-dopa decarboxylation and dopamine dehydroxylation indicated that *E. faecalis* TyrDC decarboxylates L-dopa to give dopamine (see Chapter 2) and that *E. lenta* dehydroxylates dopamine to produce *m*-tyramine using the molybdenum-dependent enzyme Dadh. Separately, these organisms and enzymes perform the two steps in the major pathway for gut microbial L-dopa metabolism. Thus, asked whether they would be sufficient for reconstitution of the full pathway in vitro. Consistent with our prediction, wild-type *E. faecalis* grown with *E. lenta* A2 (Arg506) fully converted L-dopa to *m*-tyramine (Figure 3.26). A co-culture containing the *E. faecalis* tyrDC mutant could not consume L-dopa. However, *m*-tyramine was produced when exogenous dopamine was added to this culture, revealing that *E. lenta* A2 was still metabolically active. Similarly, the gut microbial decarboxylase inhibitor AFMT potently inhibited the transformation of L-dopa into *m*-tyramine, consistent with its inhibition of *E. faecalis* TyrDC (Figure 3.27). Finally, incubating wild-type *E. faecalis* with the non-metabolizing *E. lenta* DSM2243 (Ser506) strain produced only dopamine, indicating that this Dadh variant is also inactive in a co-culture setting (Figure 3.26).

To confirm that these strains could also metabolize L-dopa in complex gut microbiotas, we added *E. faecalis* to a subset of human fecal samples that did not decarboxylate L-dopa (described in Chapter 2, Section 2.2.5). While introducing the tyrDC-deficient strain did not change L-dopa levels, including the wild-type strain led to complete depletion of L-dopa (Figure 3.28). In some samples, addition of wild-type *E. faecalis* was sufficient to yield quantitative production of *m*-tyramine, indicating the presence of dopamine dehydroxylating organisms in these communities (Figure 3.28). Finally, addition of both the wild-type *E. faecalis* and the metabolizing strain *E.
lenta A2 to non-metabolizing samples or the addition of *E. lenta* A2 alone to a decarboxylating sample generated *m*-tyramine (Figure 3.28). Taken together, these data indicate that metabolizing *E. faecalis* and *E. lenta* are sufficient for transforming L-dopa into *m*-tyramine in minimal and complex communities.

**Figure 3.26 E. faecalis and E. lenta produce m-tyramine from L-dopa in co-culture.** Metabolism of L-dopa by co-cultures of *E. faecalis* and *E. lenta* strains co-cultured for 48 hours with 1 mM *d*3-phenyl-L-dopa or 1 mM dopamine in BHI containing 0.75% arginine (w/v). Results are mean ± the SEM (n=3 replicates).
Figure 3.27 Metabolism of L-dopa by co-cultures of *E. faecalis* and *E. lenta* strains in the presence of AFMT.

Strains were co-cultured for 48 hours in BHI medium containing 0.5% arginine and 1 mM 3-phenyl-L-dopa with or without 125 μM AFMT. Metabolites in culture supernatants were analyzed by LC-MS/MS. Stacked bar plots represent the mean ± the SEM of three biological replicates.
Figure 3.28 Gain-of-function studies in stool samples incubated with \(d_3\)-phenyl-L-dopa.

A subset of the non-metabolizing stool samples from Chapter 2, Section 2.2.5 were incubated anaerobically in MEGA medium containing \(d_3\)-phenyl-L-dopa (1 mM) for 72 hours at 37 °C. LC-MS/MS was used to quantify metabolites. The samples were incubated either without any additional strains, or with \(E. faecalis\) MMH594 wild-type, \(E. faecalis\) tyrDC mutant, or \(E. lenta\) A2. Each panel represents one individual. Bars represent the mean ± the SEM of three biological replicates.
3.2.10 Carbidopa and AFMT do not impact *E. lenta* growth or dopamine dehydroxylation

As detailed in Chapter 2, L-dopa is often co-prescribed with host AADC inhibitors that limit peripheral decarboxylation. We found that carbidopa was a poor inhibitor of L-dopa decarboxylation by *E. faecalis* TyrDC while AFMT selectively and potently inhibited this activity. Having established the roles of *E. faecalis* and *E. lenta* in the cooperative metabolism of L-dopa, we wondered whether carbidopa and AFMT might impact dopamine dehydroxylation. Incubating these inhibitors individually with anaerobic cultures of *E. lenta* A2, we found that they did not inhibit growth by *E. lenta* A2 or dopamine metabolism (Figures 3.24 and 3.25). Together its lack of growth inhibition of *E. faecalis* MMH594, this finding raises the possibility that the inhibitor AFMT may be non-toxic to bacterial cells in general, although a broader screen of common gut microbes is needed to validate this claim.

**Figure 3.29** Anaerobic growth of *E. lenta* A2 in the presence of carbidopa or AFMT. *E. lenta* was grown anaerobically in BHI medium at 37 °C with varying concentrations of A) carbidopa or B) AFMT. There was no obvious dose-dependent growth defect of *E. lenta* in the presence of these inhibitors. The data shown is the mean of three replicate growth experiments ± the SEM.
Figure 3.30 Dehydroxylation of dopamine by *E. lenta* A2 in the presence of carbidopa or AFMT.

*E. lenta* A2 was grown anaerobically in BHI medium at 37 °C for 48 hours with 500 µM dopamine and varying concentrations of A) carbidopa or B) AFMT. Culture supernatants were analyzed for dopamine and *m*-tyramine by LC-MS/MS. % Dopamine dehydroxylation was calculated as the concentration of *m*-tyramine relative to the total concentration of dopamine and *m*-tyramine. There was no obvious dose-dependent defect in dopamine dehydroxylation by *E. lenta* A2 in the presence of these inhibitors. The data shown is the mean of three replicate experiments ± the SEM.
3.2.11 Attempts to reconstitute gut bacterial metabolism of L-dopa to \textit{m-}tyramine in gnotobiotic mice

We next attempted to reconstitute the full gut bacterial metabolic pathway from L-dopa to \textit{m-}tyramine in gnotobiotic mice. Dr. Elizabeth Bess (Turnbaugh lab, UCSF), colonized groups of gnotobiotic mice with four pairs of bacteria, mirroring our co-culturing setup: (1) \textit{E. faecalis} MMH594 WT and \textit{E. lenta} A2, (2) \textit{E. faecalis tyrDC} mutant and \textit{E. lenta} A2, (3) \textit{E. faecalis tyrDC} mutant and non-metabolizing strain \textit{E. lenta} DSM2243, and (4) \textit{E. faecalis} MMH594 WT and non-metabolizing strain \textit{E. lenta} DSM2243. Using qPCR assays of stool samples from these mice, I found that \textit{E. faecalis} and \textit{E. lenta} colonized the mice at similar levels across all groups (Figure 3.32). Dr. Bess then dosed the groups with $d_3$-phenyl-L-dopa (10 mg/kg) and carbidopa (30 mg/kg) and singly housed the mice in metabolic cages for 18 hours, at which time she collected serum from cardiac puncture, as well as ileal (I8), cecal, and proximal colon tissues and their intestinal contents. She also collected brains. All samples were harvested and flash-frozen in liquid nitrogen. I then extracted all tissues and used LC-MS to detect the major microbial and host-derived L-dopa metabolites, which would be expected to carry three deuterium atoms (Figure 3.31).
Figure 3.31 Predicted metabolites in gnotobiotic mice colonized with *E. faecalis* and *E. lenta* and dosed with *d*3-*phenyl-L-dopa.
Host enzymes are displayed in blue. AADC = aromatic amino acid decarboxylase. MAO = monoamine oxidase; COMT = catechol-O-methyltransferase. Microbial enzymes are displayed in red. TyrDC = tyrosine decarboxylase. Dadh = dopamine dehydroxylase.

Figure 3.32 Levels of *E. faecalis* and *E. lenta* in gnotobiotic mice.
gDNA was extracted from stool and was subjected to qPCR with *E. lenta* or *E. faecalis*-specific primers targeting the 16s rRNA gene. Stool was collected right before dosing of L-dopa and carbidopa. There was no clear difference in colonization of *E. faecalis* (A) or *E. lenta* (B) across the groups. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
We focused our initial analysis on urine, the biofluid where L-dopa metabolites are typically found (29, 80, 81). All groups displayed similar recovery of urine in the metabolic cage (Figure 3.33). There was no difference in bulk catechol excretion, as assessed by the colorimetric method for catechol detection (Figure 3.33). Because urinary L-dopa metabolites are usually conjugated to either sulfate or glucuronide groups (29, 80, 81), we first incubated urine with glucuronidase and sulfatase enzymes (from the snail *Helix pomatia*) prior to extraction. We detected significant amounts of urinary L-dopa, suggesting a large fraction is not metabolized in our experimental setup (Figure 3.34). There was no obvious difference in labeled L-dopa, dopamine, DOPAC, or homovanillic acid (HVA) between the groups (Figure 3.34). However, we detected high levels of p-tyramine, a metabolite resulting from decarboxylation of dietary or endogenous tyrosine by TyrDC, indicating that this enzyme is active in this gnotobiotic setup (Figure 3.34). We did not detect the predicted microbial metabolite m-hpaa despite recovery of the internal standard. Notably, we failed to detect the added m-tyramine internal standards, suggesting this metabolite was destroyed during our extraction. Thus, we reanalyzed these samples, substituting the enzyme-mediated hydrolysis with acid hydrolysis. Using this method, we found potential differences in levels of labeled dopamine in urine (Figure 3.35). We also observed striking differences in urinary p-tyramine between groups colonized with WT and mutant *E. faecalis*, further confirming that the TyrDC enzyme is active (Figure 3.35). Even though we detected the m-tyramine standard, indicating this metabolite survived acid hydrolysis, we did not detect labeled m-tyramine. Additionally, internal standards of *m*-HPPA, L-dopa, DOPAC, and HVA were not recovered, suggesting they were consumed during acid hydrolysis.
Figure 3.33 Urinary recovery and total catechol excretion in gnotobiotic experiment.
A) Total volume of urine collected in metabolic cages. There was no difference in the overall recovery between the groups. B) Total catechol excretion in urine, as assessed by the colorimetric assay for catechol detection. There was no difference in the overall catechol excretion between the groups. The absorbance at 500 nm is proportional to the catechol concentration. Control urine represents a mouse not dosed with L-dopa/carbidopa. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
Figure 3.34 Metabolites detected in gnotobiotic mouse urine subjected to enzymatic hydrolysis prior to extraction.

Urine was treated with *H. pomatia* sulfatase and glucuronidase prior to extraction. Each panel represents a different metabolite detected using LC-MS/MS, as indicated above the bar graphs. The parentheses show the mass transition of the metabolite. The internal standard of non-labeled *m*-tyramine was not detected following enzyme treatment, suggesting it was lost during this step. The changes in abundance of *p*-tyramine suggests *E. faecalis* TyrDC decarboxylated *p*-tyrosine in this experiment. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
Figure 3.35 Metabolites detected in gnotobiotic mouse urine subjected to acid hydrolysis prior to extraction.

Urine was subjected to acid hydrolysis prior to extraction. Bars represent mean and standard error. Each panel represents a different metabolite detected using LC-MS/MS, as indicated above the bar graphs. The parenthesis shows the mass transition of the metabolite. Labeled m-tyramine was not detected in this setup despite the internal non-labeled m-tyramine standard was detected. The only other internal non-labeled standard detected was dopamine, and we also detected labeled dopamine in urine. The other internal standards were ruined by acid treatment, suggesting that their corresponding deuterium labeled analytes also disappeared. The presence of p-tyramine suggests *E. faecalis* TyrDC decarboxylated p-tyrosine in this setup. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
We next analyzed other tissues and biological samples. In cecal contents, we detected \( p \)-tyramine but no deuterium-labeled metabolites (Figure 3.36). In stool samples subjected to enzymatic hydrolysis before extraction, we only detected deuterium labeled dopamine and unlabeled \( p \)-tyramine but no other labeled metabolites (Figure 3.37). However, because the enzymatic treatment destroyed the \( m \)-tyramine standard, we also extracted stool without any hydrolysis. In this setup, we only detected labeled L-dopa and unlabeled \( p \)-tyramine derived from tyrosine decarboxylation (Figure 3.38). In ileal contents, we only detected \( p \)-tyramine, suggesting the microbial decarboxylase is active in the small intestine, the main site of L-dopa decarboxylation (Figure 3.39). Finally, in cardiac puncture serum, we did not detect any labeled metabolites or \( p \)-tyramine despite good recoveries of standard (data not shown). In summary, across all tissues and samples, we never detected the expected \textit{meta}-hydroxyl metabolites that had been previously observed in rats \((29, 80-82)\).
Figure 3.36 Metabolites detected in gnotobiotic mouse cecal contents subjected to enzymatic hydrolysis prior to extraction.

Cecal contents were subjected to treatment with *H. pomatia* sulfatase and glucuronidase prior to extraction. The panel represents a metabolite detected using LC-MS/MS. The parenthesis shows the mass transition of the metabolite. The internal standard of non-labeled *m*-tyramine was not detected following enzyme treatment, suggesting it disappeared during enzyme treatment. Additionally, no deuterium-labeled metabolites were detected despite recovery of all other internal standards. We only detected *p*-tyramine, the presence of which suggests *E. faecalis* TyrDC decarboxylated *p*-tyrosine in this setup. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
Figure 3.37 Metabolites detected in gnotobiotic mouse stool samples subjected to enzymatic hydrolysis prior to extraction.

Stool samples were subjected to treatment with *H. pomatia* sulfatase and glucuronidase prior to extraction. Each panel represents a different metabolite detected using LC-MS/MS, as indicated above the bar graphs. The parenthesis shows the mass transition of the metabolite. The internal standard of non-labeled *m*-tyramine was not detected following enzyme treatment, suggesting it disappeared during enzyme treatment. Additionally, no deuterium-labeled metabolites other than dopamine were detected despite recovery of all other internal standards. The presence of *p*-tyramine suggests that *E. faecalis* TyrDC decarboxylated *p*-tyrosine in this setup. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
Figure 3.38 Metabolites detected in gnotobiotic mouse stool samples not subjected to enzymatic hydrolysis prior to extraction.

Stool samples were extracted without any form of hydrolysis. Bars represent mean and standard error. Each panel represents a different metabolite detected using LC-MS/MS, as indicated above the bar graphs. The parenthesis shows the mass transition of the metabolite. The only deuterium-labeled metabolite detected was L-dopa, despite good recovery of all other standards. The presence of p-tyramine suggests that E. faecalis TyrDC decarboxylated p-tyrosine in this setup. Each dot represents an individual mouse. Bars represent the mean ± the SEM.

Figure 3.39 Metabolites detected in gnotobiotic mouse ileal contents.

Ileal contents were extracted without any form of hydrolysis. Bars represent mean and standard error. The parenthesis shows the mass transition of the metabolite, which was detected using LC-MS/MS. The metabolite detected was p-tyramine, suggesting that the E. faecalis TyrDC was active in the small intestine of the gnotobiotic mice. Each dot represents an individual mouse. Bars represent the mean ± the SEM.
Our failure to detect microbial metabolites could be due to a number of reasons. For example, we utilized mice whereas all previous studies of gut microbial L-dopa metabolism have employed rats \((29, 80, 81, 83, 84)\). Drug pharmacokinetic profiles can differ between rats and mice \((85, 86)\). Although such species differences have not been explored in L-dopa metabolism, it is worth considering for future studies. In addition, we co-dosed the mice with carbidopa using a 3:1 ratio of carbidopa to L-dopa. Carbidopa was not used in previous studies of gut microbial L-dopa metabolism \((29, 80, 81, 83, 84)\). In addition, the ratio given to patients is 4:1 L-dopa to carbidopa \((87)\), suggesting that our dose may not be clinically relevant. Even though we had found that carbidopa does not impact growth or metabolism of \textit{E. faecalis} and \textit{E. lenta} in vitro, it is possible that carbidopa has unanticipated effects in vivo that could interfere with metabolism. Moreover, while the high levels of \(p\)-tyramine detected in the mice suggest TyrDC is active in our gnotobiotic experiment, it is unclear whether this enzyme also acted on L-dopa in this setup. Similarly, we did not dose mice with dopamine to verify that Dadh is also active and that dopamine dehydroxylation can occur at all under these conditions. Finally, in some cases we failed to recover the unlabeled standards, suggesting that further optimization of metabolite isolation procedures and detection is needed. Overall, further experiments are necessary to determine the factors underlying a lack of obvious microbial metabolism in these gnotobiotic mice.
3.3 Discussion

In this study, we identified gut microbial strains and enzymes that dehydroxylate dopamine, the key intermediate in the two-step gut microbial degradation of the Parkinson’s drug L-dopa and a human neurotransmitter present in the gastrointestinal tract. We find that the molybdenum-dependent enzyme Dadh dehydroxylates dopamine and that a single amino acid substitution in this enzyme correlates with metabolism by isolated strains and complex communities. We also demonstrate that *E. faecalis* and *E. lenta* encoding active Dadh cooperate to transform L-dopa into *m*-tyramine in co-cultures and complex gut microbial communities. The decarboxylase inhibitor AFMT inhibits this reaction in co-cultures without impacting growth or metabolism of *E. lenta*, consistent with our observations in complex fecal samples (described in Chapter 2). Finally, we attempt to reconstitute this pathway in gnotobiotic mice, but our failure to detect labeled *meta*-hydroxylated microbial L-dopa metabolites – even in cases where unlabeled standards were recovered – suggest further work is necessary to understand the factors that enable this microbial metabolism in vivo. It will also be important to optimize metabolite extraction and detection as we in some cases failed to recover the internal standards.

As mentioned in Chapter 2, a recent study by van Kessel and co-workers (88) characterized organisms and enzymes involved in gut microbial L-dopa decarboxylation, but their study did not explore dopamine dehydroxylation. It is possible that this uniquely microbial transformation influences side effects of L-dopa administration linked to dopamine production, and our discoveries will enable further study of this possibility. Knowledge of gut microbial genes and enzymes involved in dopamine-dehydroxylation should also enable further investigations of *m*-tyramine, whose precise targets and bioactivity in the host are unclear.
In many cases, microbial drug metabolism involves only one chemical transformation, suggesting that a single type of enzyme might be involved (89). While dietary molecules often undergo multi-step metabolism involving a diversity of microbes (60, 90), our study indicates that inter-species microbial metabolism is also relevant to drugs. Our data illustrates patients can vary in their ability to perform distinct steps of a metabolic pathway, suggesting that the full diversity of reactions that a drug undergoes must be considered to understand the effects of microbial metabolism on drug efficacy and patient response. For example, while microbial L-dopa decarboxylation would remove this drug from circulation, the second step of the pathway, dopamine dehydroxylation, could minimize side effects of peripheral dopamine and produce additional biologically active metabolites. It is possible that variations in each step of the L-dopa pathway results in variable drug responses.

In addition to highlighting microbial metabolism as a community effort, our study uncovers an unexpected role of SNPs in the gut microbiota. The influence of SNPs is well-established in the context of the human genome. SNPs in drug transporters, xenobiotic metabolizing enzymes, or regulatory regions influence human metabolism of a number of drugs, including cancer treatments, anti-virals, anti-depressants, and antihypertensive drugs (91-94). More recent work has revealed potential roles for SNPs in gut microbial metabolism. For instance, a SNP in the *treR* transcriptional regulator has been implicated in the growth of the intestinal pathogen *Clostridium difficile* on the food additive trehalose (79). Similarly, a SNP in the *E. lenta* Cgr1 enzyme correlates with digoxin reduction by purified enzyme and isolated strains, with a tyrosine variant displaying significantly higher turnover rates than an asparagine variant in vitro (57). To our knowledge, our work is the first to report a link between a SNP and metabolism by complex gut microbiota communities. This finding reinforces the idea that microbial phylogeny very often does not predict
function. However, our results suggest that even simply detecting functional genes may not accurately predict the activities encoded by the human gut microbiome, underscoring the importance of studying and characterizing enzymes from this community.

Additionally, our study is a reminder that strains found at low abundance perform important functions in the human gut microbiota. Researchers in the human gut microbiota field sometimes equate gut microbial abundance with microbial function and importance in the host. The work described in this chapter adds to a growing body of evidence challenging this idea. Although *E. lenta* and its close relatives within the in the Coriobacteriia class are generally found at low abundances in the healthy human gut microbiota, these strains are emerging as key players in microbial metabolism. These gut bacteria catalyze wide range of clinically relevant biotransformations, including the reduction of the cardiac drug digoxin, oxidation of host-derived bile acids, and reduction of prominent dietary polyphenols (57, 61, 62, 95, 96). These strains are also considered opportunistic pathogens and have been associated with autoimmune diseases (97, 98). Our study adds dopamine dehydroxylation to a growing list of important Coriobacteriia functions and highlights that this relatively neglected group of bacteria might be far more important in the human gut microbiota than previously appreciated. Further studies on the fundamental biology of *E. lenta* and its close relatives hold promise to uncover new metabolic transformations and provide new insights into the inner workings of both microbiota and the host.

Our study will inform future work on neurotransmitter metabolism and catechol dehydroxylation in the human gut. Our findings support the idea that gut microbes can catabolize host-derived neurotransmitters (99) and raise questions about the biological consequences of gut microbial metabolism of endogenous dopamine. Dopamine is produced at substantial levels in the gastrointestinal tract and is available in the gut lumen (6, 13). Consequently, our findings raise
questions about how microbial dehydroxylation affects known dopamine-dependent processes such as gut motility and the colonization and virulence of intestinal pathogens (1, 6, 21). Finally, catechol dehydroxylation, the type of chemistry involved in dopamine dehydroxylation, is a widespread microbial metabolic reaction involved in transforming a variety of drugs and diverse dietary compounds in the human gut microbiota (84, 100-104). Further studies of the Dadh substrate scope and transcriptional regulation will illuminate whether Dadh is the main enzyme responsible for catechol dehydroxylation in the human gut microbiota or whether this enigmatic microbial reaction is performed by more specialized enzymes. Efforts to understand the molecular basis catechol dehydroxylation and its physiological relevance to gut microbes will be described in Chapter 4 of this thesis.
3.4 Methods

3.4.1 General materials

The following chemicals were used in this study: L-dopa (Sigma-Aldrich, catalog# D9628-5G), dopamine (Sigma-Aldrich, catalog# PHR1090-1G), m-tyramine (Santa Cruz Biotechnology, catalog# sc-255257), Isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich, catalog# I5502-1G), carbidopa (Sigma-Aldrich, catalog# PHR1655-1G), L-arginine (Sigma-Aldrich, catalog# A5006-100G), sodium molybdate (Sigma-Aldrich, catalog # 243655-100G), sodium tungstate (72069-25G), SIGMAFAST protease inhibitor tablets (Sigma-Aldrich, catalog#: S8830), (S)-α-fluoromethyltyrosine (AFMT) (obtained from Merck Sharp & Dohme Corp under MTA LKR166502), d3-phenyl-L-dopa (Sigma-Aldrich, catalog# 333786-250MG), benzyl viologen (Sigma-Aldrich, catalog# 271845-250mg), methyl viologen (Sigma-Aldrich, catalog# 856177-1g), diquat (Sigma-Aldrich, catalog# 45422-250mg), sodium dithionite (Sigma-Aldrich, catalog# 157953-5G), p-tyramine (Sigma-Aldrich, catalog# T90344-5G), DOPAC (Sigma-Aldrich, catalog# 850217-1G), Homovanillic acid (Sigma-Aldrich, catalog# H1252-100MG), 3-hydroxyphenylacetic acid (Sigma-Aldrich, catalog# H49901-5G). Acetonitrile and methanol for LC-MS analyses were purchased as LC-MS grade solvent from Honeywell Burdick & Jackson or Sigma-Aldrich.

3.4.2 General methods

All bacterial culturing work was performed in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen, 10% carbon dioxide, and nitrogen as the balance. Hungate tubes were used for anaerobic culturing unless otherwise noted (Chemglass, catalog# CLS-4209-01). All lysate work and biochemical experiments were performed in an anaerobic
chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen and nitrogen as the balance situated in a cold room at 4 °C.

All genomic DNA (gDNA) was extracted from bacterial cultures using the DNeasy UltraClean Microbial Kit (Qiagen, catalog #: 12224-50) according to the manufacturer’s protocol.

3.4.3 LC-MS methods

Method A: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire PFP column (4.6 × 100 mm, 3.5 μm; catalog #81601). The flow rate was 1.0 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 0% B isocratic, 2-9 min: 0-10% B, 9-11 min: 10-95% B, 11-13 min: 95% B isocratic, 13-15 min: 95-0% B, 15-18 min: 0% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of \( d_3 \)-phenyl-L-dopa (precursor ion m/z = 201.3, daughter ion m/z = 155.2), L-dopa (precursor ion m/z = 198.3, daughter ion m/z = 152.2), \( d_3 \)-phenyl-dopamine (precursor ion m/z = 157.3, daughter ion m/z = 140.3), dopamine (precursor ion m/z = 154.3, daughter ion m/z = 137.3), \( d_3 \)-phenyl-tyramine (precursor ion m/z = 141.3, daughter ion m/z = 124.3), tyramine (precursor ion m/z = 138.3, daughter ion m/z = 121.3) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode.

Method B: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire PFP column (4.6 × 100 mm, 3.5 μm; catalog #81601). The flow rate was 1.0 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-9 min: 0-10% B, 9-10 min: 10-30% B, 10-15 min: 30-37%
B, 15-16 min: 37-95%, 16-18 min: 95% B isocratic, 18-19.5 min: 95%-0%, 19.5-22 min: 0% B isocratic. The instrument was set to positive mode between minutes 0-11 and was set to negative mode during minutes 11-15. For mass spectrometry in positive mode:, the source temperature was 300 °C, and the masses of $d_3$-phenyl-L-dopa (precursor ion m/z = 201.3, daughter ion m/z = 155.2), L-dopa (precursor ion m/z = 198.3, daughter ion m/z = 152.2), $d_3$-phenyl-dopamine (precursor ion m/z = 157.3, daughter ion m/z = 140.3), dopamine (precursor ion m/z = 154.3, daughter ion m/z = 137.3), $d_3$-phenyl-tyramine (precursor ion m/z = 141.3, daughter ion m/z = 124.3), tyramine (precursor ion m/z = 138.3, daughter ion m/z = 121.3) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in MRM mode. For mass spectrometry in negative mode: the source temperature was 300 °C, and the masses of $d_3$-phenyl-homovanillic acid (precursor ion m/z = 184.2, daughter ion m/z = 140.2), homovanillic acid (precursor ion m/z = 181.2, daughter ion m/z = 137.2) were monitored at a collision energy of 5 mV and fragmentor setting of 135 in MRM mode. Additionally, $d_3$-phenyl-DOPAC (precursor ion m/z = 170.2, daughter ion m/z = 126.2), DOPAC (precursor ion m/z = 167.2, daughter ion m/z = 123.2), $d_3$-phenyl-hydroxyphenylacetic acid (precursor ion m/z = 154.3, daughter ion m/z = 110.3), hydroxyphenylacetic acid (precursor ion m/z = 151.3, daughter ion m/z = 121.3) were monitored at a collision energy of 3 mV and fragmentor setting of 75 in MRM mode.

Method C: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim polar advantage II (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min$^{-1}$ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. The same masses as in Method A were monitored using the same settings.
3.4.4 Screen of a collection of gut bacteria in anaerobic culture for dopamine dehydroxylation using a colorimetric assay

Our colorimetric assay for dopamine dehydroxylation was based on the Arnow test (52). Briefly, 50 µL of 0.5 M HCl was added to 50 µL of culture supernatant. After mixing, 50 µL of an aqueous solution containing both sodium molybdate and sodium nitrite (0.1 g/mL each) was added, which produced a yellow color. Finally, 50 µL of 1 M NaOH was added followed by pipetting up and down to mix. This allowed the characteristic pink color to develop. Absorbance was measured at 500 nm immediately using a Synergy HTX Multi-Mode Microplate Reader (BioTek). Initially, we grew a set of gut strains in MEGA media anaerobically for 48 hours with 250 µM dopamine and used the colorimetric assay to assess their ability to dehydroxylate dopamine. None of the strains screened (Enterococcus faecalis MMH594, Enterococcus faecalis TX0104, Enterococcus faecium TX01330, Clostridium aspargiforme DSM15981, Flavonifractor plautii ATCC29863, Clostridium sp. ATCC BAA-442, Bifidobacterium longum spp. Infantis ATCC 15697, Flavonifractor plautii ATCC 49531, Clostridium ramosum DSM 1402, Akkermansia muciniphila ATCC BAA-835, Clostridium bartlettii ATCC BAA-827, Enterococcus raffinosus ATCC 49464, Eubacterium limosum ATCC 10825, Eggerthella lenta DSM2243, Faecalibacterium prausnitzii ATCC 27766, Eubacterium rectale ATCC 33656, Lactobacillus reuteri ATCC 23272, Parabacteroides distasonis ATCC 8503) displayed any detectable metabolism.

3.4.5 Collection of fecal samples from neurologically healthy human patients

The indicated stool samples were collected from 12 neurologically healthy individuals during the control phase of an inpatient study described in detail elsewhere (105) and from 7 healthy control subjects sampled at the University of California, San Francisco (UCSF). All
subjects consented to participate in the study, which was approved by the relevant Institutional Review Boards. These were collected by the Turnbaugh lab at UCSF. Samples were stored as fecal slurries in PBS and 20% glycerol at −80 °C until use.

3.4.6 Collection of fecal samples from Parkinson’s disease patients

The indicated stool samples were obtained from the BioCollective Microbiome Stool Bank (Denver, CO), which collected the samples under the Western IRB WIRB #1164096 for human subjects research. Samples were stored as fecal slurries in PBS and 20% glycerol at −80 °C until use. Specific samples (n = 12) were selected from a larger set to minimize confounding variables, and they include 6 drug-naïve individuals and 6 patients taking L-dopa/carbidopa. Participants were not taking: antibiotics, antihistamines, laxatives, suppositories, beta blockers, statins, proton pump inhibitors, tricyclic antidepressents, selective serotonin reuptake inhibitors (SSRIs), platelet aggregators, oral contraceptives, oral metformin, and nonsteroidal anti-inflammatory drugs. They did not have additional comorbidities. Groups were balanced on gender (drug-naïve = 3 males and 3 females vs. L-dopa/carbidopa = 4 males and 2 females), age (drug-naïve = 68 ± 6 years vs. L-dopa/carbidopa = 60 ± 8 years), and body mass index (drug-naïve = 23.1±4.4 kg/m² vs L-dopa/carbidopa = 25.3±3.8 kg/m²) for drug naïve and L-dopa/carbidopa patients respectively (mean ± standard deviation (SD)).

3.4.7 Enrichment culturing and isolation of a dopamine dehydroxylating gut bacterial strain

To prepare the growth medium used for enrichment culturing, 10 g NaCl, 5 g MgCl₂*6H₂O, 2 g KH₂PO₄, 3 g NH₄Cl, 3 g KCl, 0.15 g CaCl₂ x 2H₂O, 1 g yeast extract, 1 g tryptone, 10 mL of Trace Mineral Supplement (ATCC, catalog# MD-TMS), and 0.25 mL of 0.1% resazurin (dissolved in MilliQ water) were added to a final volume of 1 L of water. This medium was then boiled with stirring to dissolve all components. While cooling, the medium was bubbled with argon to ensure
it was anaerobic. Once the medium was cool, 30 mM NaHCO$_3$ and 0.4 mM L-cysteine HCl were added. The medium was then brought into an anaerobic chamber and was distributed in 10 mL aliquots into individual Hungate tubes. These tubes were autoclaved and brought back into an anaerobic chamber where 100 µL of Vitamin Supplement (ATCC, catalog# MD-VS) was added to each tube. The tubes were stored at 4 ºC.

During the enrichment culturing, all experiments were performed under strictly anaerobic conditions in an anaerobic chamber. First, a stool sample from a healthy human donor was resuspended in pre-reduced PBS at a final concentration of 0.1 g/mL. The mixture was vortexed to produce a homogenous slurry and was then left for 30 minutes to let particulates settle. The supernatant was diluted 1:10 in PBS and was further diluted 1:100 into the medium described above containing 500 µM dopamine as the electron acceptor and 10 mM sodium acetate as the electron donor. These initial cultures were incubated at 37 ºC for five days and were then passaged twice by a 1:100 dilution into fresh medium. Each successive passage was incubated for 48 hours. The final enrichment culture was streaked onto agar plates containing the basal medium described above. Once individual colonies appeared, they were picked and inoculated into the same liquid medium and the liquid cultures were screened for dopamine dehydroxylation after 48 hours of growth at 37 ºC using the colorimetric assay described above. Colonies that displayed activity in the minimal medium were diluted 1:10 in PBS and plated onto MEGA Medium agar plates to support robust growth (106) at 37 ºC for five days. Colonies that appeared on the MEGA medium plates were then inoculated into liquid MEGA medium, and metabolically active colonies were re-streaked twice on MEGA Medium agar plates to obtain pure cultures. Throughout the experiment, dopamine dehydroxylation was tracked using the colorimetric assay for catechol detection. In addition, gDNA was harvested from cultures at different stages of enrichment using the DNeasy
UltraClean Microbial Kit, allowing for 16S rRNA gene sequencing to be performed.

Sequencing to determine microbial community composition was performed by the Microbial Omics Core (MOC) at the Broad Institute. 16S rRNA gene libraries targeting the V4 region of the 16S rRNA gene were prepared by first normalizing template concentrations and determining optimal cycle number by way of qPCR. Two 25 μL reactions for each sample were amplified with 0.5 units of Phusion polymerase with 1X High Fidelity buffer, 200 μM of each dNTP, 0.3 μM of 515F (5’-AATGATACGGCGACCACCGAGATCTACACTATGGTAATTGTGTGCCAGCMGCCGCGGTAA-3’) and 806rcbc0 (5’-CAAGCAGAAGACGGCATACGAGATATCCCTTGTCTCCAGTCAGTCAGCCGGACTACHVGGGTWTCTAAT-3’). 0.25 μL of 100x SYBR were added to each reaction mixture, and samples were quantified using the formula 1.75^(deltaCt). To ensure minimal over-amplification, each sample was diluted to the lowest concentration sample, amplifying with this sample optimal cycle number for the library construction PCR. Four 25 μL reactions were prepared per sample with master mix conditions listed above, without SYBR. Each sample was given a unique reverse barcode primer from the Golay primer set (107, 108). Replicates were then pooled and cleaned via Agencourt AMPure XP-PCR purification system. Purified libraries were diluted 1:100 and quantified again via qPCR (Two 25 μL reactions, 2x iQ SYBR SUPERMix (Bio-Rad, REF: 1708880 with Read 1 (5’-TATGGTAATT GT GTGYCAGCMGCCGCGGTAA-3’), Read 2 (5’-AGTCAGTCAG CC GGACTACNVGGGTWTCTAAT-3’)). Undiluted samples were normalized by way of pooling using the formula mentioned above. Pools were quantified by Qubit (Life Technologies, Inc.) and normalized into a final pool by Qubit concentration and number of samples. Final pools were sequenced on an Illumina MiSeq 300 using custom index 5’-
ATTAGAWACCCBDGTAGTCCGGCTGACTGACT-3’ and custom Read 1 and Read 2 mentioned above. Sequencing data were analyzed using QIIME (109).

We identified *Eggerthella lenta* A2 as an active dopamine dehydroxylating strain. Following extraction of the gDNA, we sequenced the genome using PacBio and Illumina. The Illumina sequencing and subsequent genome annotation was performed by the Turnbaugh lab at UCSF. The PacBio sequencing and subsequent genome annotation was performed by Era7 Bioinformatics, Inc. The details are described in a paper by our collaborators in the Turnbaugh lab at UCSF. This paper describes the collection and sequencing of the Actinobacterial collection described in this chapter (77).

### 3.4.8 Culture-based assays to assess the inducibility of dopamine dehydroxylation in *E. lenta* A2

10 mL turbid 48-hour starter cultures of *E. lenta* A2 in BHI medium were diluted 1:100 into 10 mL of BHI medium with or without 500 μM dopamine. In mid-exponential phase (approximately OD$_{600}$=0.060) cells were pelleted by centrifugation following by resuspension into 10 mL PBS to wash the cells. Cells were pelleted again by centrifugation and were redissolved in BHI medium containing either vehicle (water) or dopamine (500 μM final concentration). Culture aliquots were collected at regular intervals and were harvested by centrifugation. Supernatants were analyzed for metabolism using the colorimetric assay for catechol detection.

### 3.4.9 Lysate assays to assess dopamine dehydroxylation in *E. lenta* A2

10 mL turbid 48-hour starter cultures of *E. lenta* A2 in BHI medium were diluted 1:100 into 20 mL BHI medium containing 1% arginine and 10 mM formate with or without 500 μM dopamine. Cells were pelleted by centrifugation when cultures reached OD$_{600}$=0.500. Cell pellets were washed twice with PBS and were then re-suspended in 800 μL 50 mM Tris pH 8 containing
4 mg/mL SIGMAFAST protease inhibitor cocktail, followed by anaerobic sonication to lyse the cells. Dopamine was added to crude lysates at a final concentration of 500 µM and the reactions were left for 12 hours at room temperature under anaerobic conditions. To assess the impact of oxygen on dopamine dehydroxylation by cell lysates, the assay was also set up outside the anaerobic chamber. To assess the impact of tungstate on dopamine dehydroxylation by cell lysates, sodium tungstate (2.5, 5, and 10 mM) was added at the time of dopamine addition. After incubation, all lysates were dissolved 1:10 with LC-MS grade methanol and centrifuged to pellet precipitates. Supernatant were analyzed by LC-MS/MS to measure dopamine and \( m \)-tyramine production using Method A described above. Experiments were performed in triplicate and repeated twice to ensure consistency.

3.4.10 Evaluation of the impact of arginine and formate on growth and metabolism by \( E. lenta \)

\( E. lenta \) was grown for 48 hours in BHI medium anaerobically at 37 °C to generate turbid starter cultures. For assays assessing the impact of arginine on dopamine dehydroxylation when the dopamine was added at the time of inoculation, the strain was then diluted 1:100 into 200 µL BHI medium containing varying concentrations of arginine and 500 µM dopamine. Cells were grown anaerobically at 37 °C for 48 hours, and cultures were grown in 96-well plates (VWR, catalog# 29442-054). Growth was monitored using a Synergy HTX Multi-Mode Microplate Reader (BioTek) by measuring absorbance at 600 nm. Metabolism was measured using the colorimetric assay for catechol detection. For assays assessing impact of arginine on dopamine dehydroxylation when the dopamine was added during exponential growth, the starter cultures were diluted 1:100 into 5 mL BHI medium containing either no arginine, 1% arginine, or 1% arginine. Dopamine was added (500 µM final concentration) at approximately OD600 = 0.200.
Culture aliquots were collected at regular intervals and were harvested by centrifugation. Supernatants were diluted 1:10 in LC-MS grade methanol, followed by centrifugation to pellet precipitates. Metabolites were detected on LC-MS using Method A described above.

3.4.11 Screen of gut Actinobacterial isolates for dopamine dehydroxylation

Cells were cultured in 96-well plates and all experiments were performed anaerobically. The strains screened for dopamine dehydroxylation have been previously described (77). Individual strains were plated from glycerol stocks onto BHI agar plates containing 1% arginine and grown for 3 days. Colonies were then inoculated into BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures, which were diluted 1:100 in triplicate into 200 µL fresh BHI medium containing 500 µM dopamine. These cultures were grown for 48 hours at 37 °C. Cultures were harvested by centrifugation and the supernatant was diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS using Method C described above. The experiment was repeated twice.

3.4.12 RNA-sequencing in *E. lenta* A2, *E. lenta* DSM2243, and *E. lenta* 28b

Turbid 48-hour starter cultures of *E. lenta* in BHI medium were inoculated 1:100 in triplicate or quadruplicate into 5 mL of BHI medium containing 1% arginine and 10 mM formate, and grown at 37 °C anaerobically. When the cultures reached OD₆₀₀=0.200, dopamine or vehicle (water) was added at a final concentration of 500 µM. Cultures were harvested when they reached OD₆₀₀=0.500. They were centrifuged, and cell pellets were resuspended in 500 µL Trizol reagent (ThermoFisher, catalog#: 15596026) followed by flash-freezing in liquid nitrogen.

All work following freezing of bacterial pellets was performed by the MOC at the Broad Institute. Total RNA was isolated first using bead beating to lyse cells and then using the Zymo Research Direct-Zol RNA MiniPrep Plus kit (Catalog # R2070) according to the manufacturer’s
protocol. Illumina cDNA libraries were generated using a modified version of the RNAtag-Seq protocol as described (110). Briefly, 500 ng of total RNA was fragmented, depleted of genomic DNA, and dephosphorylated prior to its ligation to DNA adapters carrying 5’-AN8-3’ barcodes with a 5’ phosphate and a 3’ blocking group. Barcoded RNAs were pooled and depleted of rRNA using the RiboZero rRNA depletion kit (Epicentre). These pools of barcoded RNAs were converted to Illumina cDNA libraries in 3 main steps: (i) reverse transcription of the RNA using a primer designed to the constant region of the barcoded adaptor; (ii) addition of a second adapter on the 3’ end of the cDNA during reverse transcription using SmartScribe RT (Clonetech) as described (111); (iii) PCR amplification using primers that target the constant regions of the 3’ and 5’ ligated adaptors and contain the full sequence of the Illumina sequencing adaptors. cDNA libraries were sequenced on Illumina HiSeq 2500. For the analysis of RNAtag-Seq data, reads from each sample in the pool were identified based on their associated barcode using custom scripts, and up to 1 mismatch in the barcode was allowed with the caveat that it did not enable assignment to more than one barcode. Barcode sequences were removed from reads, and the reads were mapped to the genome of the GenBank assembly of E. lenta A2 (Genome ID PPUL00000000) with Bowtie2 (112) and feature counts derived using Rsubread (113). Differential expression analysis was carried out using DESeq2 (114) with a significant threshold of FDR<0.1 and an absolute log2 fold change of 1. For preparation of the Manhattan plot (differential expression by gene location), scaffolds were concatenated and a pseudo base position used.
3.4.13 Evaluation of the effects of tungstate and molybdate on dopamine dehydroxylation in cultures of *E. lenta* A2

Starter cultures of *E. lenta* A2 were grown over 48 hours in 10 mL of BHI medium and then inoculated 1:100 into 200 µL of BHI medium containing 500 µM dopamine and either sodium tungstate (250 µM), sodium molybdate (1000 µM), or vehicle (water). Cultures were grown for 48 hours anaerobically at 37 °C and were harvested by centrifugation. Supernatants were dissolved 1:10 in LC-MS grade methanol and analyzed using LC-MS/MS Method A described above. Experiments were performed anaerobically, and cultures were grown in 96-well plates.

3.4.14 Distribution of *E. lenta* A2 *dadh* among sequenced genomes

The *E. lenta* A2 dopamine dehydroxylase (*Dadh*) protein sequence (accession #WP_086414988.1) was used as the query sequence for a BLASTP search of the NCBI non-redundant protein database (May 1, 2018). The 500 highest-scoring sequences were exported, but only the top sequences (down to ~80% ID/E=0 score) were considered likely hits.

3.4.15 Assessment of the conservation of *dadh* and surrounding genes across a gut Actinobacterial library

Reads from genome sequencing were mapped to the reference A2000003 contig using Bowtie2 and filtered for a minimum mapping quality of 10. Variants were called when >80% of reads supported an alternate sequence. The phylogenetic tree of *E. lenta* strains was prepared previously (57, 77). This analysis was done by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).

3.4.16 Conservation of *Eggerthella lenta* A2 *dadh* among Actinobacterial isolates and alignment of protein sequences

The *Eggerthella lenta* A2 dopamine dehydroxylase protein sequence was used as the query sequence for a BLASTP search of 26 previously sequenced gut Actinobacterial genomes (77)
The genomes were loaded in Geneious (version 11) and BLASTP hits with an amino acid ID of >80% and e-value of 0 were considered hits. These sequences were extracted and aligned using Jalview version 2.10.4, allowing for identification of the amino acid residue at position 506.

### 3.4.17 Ultracentrifugation of *E. lenta* A2 lysates to assess localization of dopamine dehydroxylase activity

All experiments were performed under strictly anaerobic conditions at 4 °C. Procedures outside the anaerobic chamber were performed in tightly sealed containers to prevent oxygen contamination. First, *E. lenta* A2 starter cultures were inoculated from single colonies into liquid BHI medium and were grown for 30 hours. Starter cultures were diluted 1:100 into 4 L of BHI medium containing 1% arginine and 10 mM formate and grown anaerobically at 37 °C for 17 hours. Cells were pelleted in 4 separate 1 L bottles by centrifugation and each pellet was resuspended in 20 mL 20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail. Resuspended cells were then lysed using sonication in an anaerobic chamber. The lysates were then clarified by centrifugation. The soluble fractions were then subjected to ultracentrifugation at 35000 rpm for 1 hour and 45 minutes to pellet membranes (115). The membrane pellet was re-suspended in lysis buffer containing 0.1% Triton X-100 (v/v) (Fisher scientific, catalog# AC215680010 to solubilize membranes. Both this fraction and the soluble fractions were subjected to activity assays. In these assays, 50 µL of fraction was mixed, in the following order, with 1 µL electron donors (final concentration 1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water), 2 µL sodium dithionite (2 mM final concentration, dissolved in water), and 1 µL dopamine (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for
12–14 hours to allow dopamine dehydroxylation to proceed, followed by dilution 1:20 into LC-MS grade methanol to stop the reaction. The diluted reactions were centrifuged to pellet any precipitates and the supernatant was analyzed by LC-MS using Method A described above.

### 3.4.18 Precipitation of *E. lenta* A2 lysates using ammonium sulfate

All experiments were performed under strictly anaerobic conditions at 4 °C. Procedures outside the anaerobic chamber were performed in tightly sealed containers to prevent oxygen contamination. First, *E. lenta* A2 starter cultures were inoculated from single colonies into liquid BHI medium and were grown for 30 hours. Starter cultures were diluted 1:100 into 4 L of BHI medium containing 1% arginine and 10 mM formate and grown anaerobically at 37 °C for 17 hours. Cells were pelleted in 4 separate 1 L bottles by centrifugation and each pellet was resuspended in 20 mL 20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail. Resuspended cells were then lysed using sonication in an anaerobic chamber. The lysates were then clarified by centrifugation and 10 mL was aliquoted into a range of different 15 mL tubes. Varying amounts of solid ammonium sulfate was added to these tubes, which were left to precipitate proteins for 1 hour and 30 minutes. The tubes were then spun down to separate precipitates form the soluble fraction. The pellets were re-suspended in 10 mL lysis buffer, and both the resuspended pellets and the soluble fraction at each concentration of ammonium sulfate were subjected to activity assays. In these assays, 50 µL of fraction was mixed, in the following order, with 1 µL electron donors (final concentration 1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water), 2 µL sodium dithionite (2 mM final concentration, dissolved in water), and 1 µL dopamine (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for 12–14 hours to allow dopamine dehydroxylation to proceed, followed by dilution 1:20 into LC-MS grade
methanol to stop the reaction. The diluted reactions were centrifuged to pellet any precipitates and the supernatant was analyzed by LC-MS using Method A described above.

3.4.19 Anaerobic activity-based purification of E. lenta A2 Dadh

Protein purification: All experiments were performed under strictly anaerobic conditions at 4 °C. Procedures outside the anaerobic chamber were performed in tightly sealed containers to prevent oxygen contamination. First, E. lenta A2 starter cultures were inoculated from single colonies into liquid BHI medium and were grown for 30 hours. Starter cultures were diluted 1:100 into 4 L of BHI medium containing 1% arginine and 10 mM formate and grown anaerobically at 37 °C for 17 hours. Cells were pelleted in 4 separate 1 L bottles by centrifugation and each pellet was resuspended in 20 mL 20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail. Resuspended cells were then lysed using sonication in an anaerobic chamber. The sonication protocol involved 25 % amplitude, 2 minutes sonication total, with 10 seconds on and 40 seconds off. The lysates were then clarified by centrifugation and the soluble fractions were subjected to two rounds of ammonium sulfate precipitation. During the precipitation, two of the four 20 mL clarified lysates were combined into a final volume of 40 mL, creating two 40 mL clarified lysates from the original 4 L culture. Solid ammonium sulfate was then dissolved in these lysates at a final concentration of 30% (w/v) and lysates were left for 1 hour and 20 minutes followed by centrifugation to pellet the precipitates. The supernatant was mixed with ammonium sulfate to achieve a final concentration of 40% (w/v) and left for 1 hour and 20 minutes. Following centrifugation, each pellet containing the precipitated proteins was re-dissolved in 20 mL 20 mM Tris pH 8 containing 0.5 M ammonium sulfate. The re-dissolved pellets were combined, and the resulting 40 mL were injected onto an FPLC (Bio-Rad BioLogic DuoFlow System equipped with GE Life Sciences DynaLoop90) for hydrophobic interaction chromatography (HIC) using 5 x 1mL
HiTrap phenyl HP columns (GE Life Sciences, catalog# 17135101). Fractions were eluted with a gradient of 0.5 M to 0 M ammonium sulfate (in 20 mM Tris pH 8) at a flow rate of 1 mL/min and were tested for activity using the assay described below. The majority of the dopamine dehydroxylase activity eluted around 0.05 M-0.1 M ammonium sulfate. Active fractions displaying >50% conversion of dopamine were combined and injected onto the FPLC system described above for anion exchange chromatography using a UNO Q1 column (Bio-Rad, catalog# 720-0001) at a flow-rate of 1 mL/min. Fractions were eluted using a gradient of 0 to 1 M NaCl in 20 mM Tris pH 8 and were tested for activity. The majority of the dopamine dehydroxylase activity eluted around 250 mM NaCl. Active fractions were combined and concentrated 20-fold using a spin concentrator with a 5 kDa cutoff. The concentrate was injected onto FPLC for size exclusion chromatography using an Enrich 24 mL column (Enrich SEC 650, 10*300 column, Bio-Rad, catalog# 780-1650). Fractions were eluted over a 26 mL volume run isocratically in 20 mM Tris pH 8 containing 250 mM NaCl, collected, and were subjected to activity assays. Active fractions were analyzed using SDS-PAGE to assess the presence of protein.

**Activity assays:** 50 µL aliquote of the fractions from FPLC were mixed in the following order with 1 µL electron donors (final concentration 1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water), 2 µL sodium dithionite (2 mM final concentration, dissolved in water), and 1 µL dopamine (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for 12–14 hours to allow dopamine dehydroxylation to proceed, followed by dilution 1:20 into LC-MS grade methanol to stop the reaction. The diluted reactions were centrifuged to pellet any precipitates and the supernatant was analyzed by LC-MS using Method A described above. All work for proteomics (preparation of samples, detection on LC-MS/MS, and data analysis) described below.
was performed by the Bauer Core at Harvard University.

Proteomics: Sample preparation, global proteomics – To 250 µL of fraction 5 shown in Figure S11 was added 3 µL of 20 mM Tris(2-carboxyethyl)phosphine (TCEP, Sigma-Aldrich, catalog# 75259) in 50 mM TEAB triethylammonium bicarbonate (TEAB, Sigma-Aldrich, catalog# T7408). The mixture was incubated at 37 °C for 1 hour in a sealed tube. The mixture was cooled to room temp for 10 minutes, followed by vortexing and centrifugation. To this mixture was added 3 µL of freshly prepared 40 mM iodoacetamide Ultra (Sigma-Aldrich, catalog# I1149-5G) in 50 mM TEAB. The reaction mixture was incubated in a sealed tube for 1 hour at room temperature under tin foil to block light. 0.5 µL of trypsin (Promega, catalog# V5111) was added, and the mixture was incubated for 16 hours at 37 °C in a thermocycler. This sample was used for protein identification by LC-MS/MS, as described below.

Sample preparation, gel band corresponding to dopamine dehydroxylase – The cut-out gel band was washed twice with 50% aqueous acetonitrile for 5 minutes followed by drying in a SpeedVac. The gel was then reduced with a volume sufficient to completely cover the gel pieces (100 µL) of 20 mM TCEP in 25 mM TEAB at 37 °C for 45 minutes. After cooling to room temp, the TCEP solution was removed and replaced with the same volume of 10 mM iodoacetamide Ultra (Sigma) in 25 mM TEAB and kept in the dark at room temperature for 45 minutes. Gel pieces were washed with 200µL of 100 mM TEAB (10 minutes). The gel pieces were then shrunk with acetonitrile. The liquid was then removed followed by swelling with the 100 mM TEAB again and dehydration/shrinking with the same volume of acetonitrile. All of the liquid was removed, and the gel was completely dried in a SpeedVac for ~20 minutes. 0.06 µg/5 µL of trypsinin 50 mM TEAB was added to the gel pieces and the mixture was placed in a thermomixer at 37°C for about 15 min. 50 µL of 50 mM TEAB was added to the gel slices. The samples were vortexed,
centrifuged, and placed back in the thermomixer overnight. Samples were digested overnight at 37 °C. Peptides were extracted with 50 µL 20 mM TEAB for 20 min and 1 change of 50 µL 5% formic acid in 50% acetonitrile at room temp for 20 minutes while in a sonicator. All extracts obtained were pooled into an HPLC vial and were dried using a SpeedVac to the desired volume (~50 µL). This sample was used for protein identification by LC-MS/MS, as described below.

**Mass spectrometry:** Each sample was used for a single LC-MS/MS experiment that was performed on an LTQ Orbitrap Elite (Thermo Fischer) equipped with a Waters (Milford, MA) NanoAcquity HPLC pump. Peptides were separated using a 100 µm inner diameter microcapillary trapping column packed first with approximately 5 cm of C18 Reprosil resin (5 µm, 100 Å, Dr. Maisch GmbH, Germany) followed by ~20 cm of Reprosil resin (1.8 µm, 200 Å, Dr. Maisch GmbH, Germany). Separation was achieved through applying a gradient of 5–27% ACN in 0.1% formic acid over 90 min at 200 nL min⁻¹. Electrospray ionization was enabled through applying a voltage of 1.8 kV using a home-made electrode junction at the end of the microcapillary column and sprayed from fused silica pico tips (New Objective, MA). The LTQ Orbitrap Elite was operated in data-dependent mode for the mass spectrometry methods. The mass spectrometry survey scan was performed in the Orbitrap in the range of 395 –1,800 m/z at a resolution of 6 × 10⁴, followed by the selection of the twenty most intense ions (TOP20) for CID-MS2 fragmentation in the Ion trap using a precursor isolation width window of 2 m/z, AGC setting of 10,000, and a maximum ion accumulation of 200 ms. Singly charged ion species were not subjected to CID fragmentation. Normalized collision energy was set to 35 V and an activation time of 10 ms. Ions in a 10 ppm m/z window around ions selected for MS2 were excluded from further selection for fragmentation for 60 seconds. The same TOP20 ions were subjected to HCD MS2 event in Orbitrap part of the instrument. The fragment ion isolation width was set to 0.7 m/z,
AGC was set to 50,000, the maximum ion time was 200 ms, normalized collision energy was set to 27 V and an activation time of 1 ms for each HCD MS2 scan.

Mass spectrometry data analysis: Raw data were submitted for analysis in Proteome Discoverer 2.1.0.81 (Thermo Scientific) software. Assignment of MS/MS spectra were performed using the Sequest HT algorithm by searching the data against a protein sequence database including a custom database from Eggerthella lenta A2 and entries from the Human Uniprot database (16,768 proteins from SwissProt and 62,460 proteins from TrEMBL for a total of 79,228 protein forms, 2015) and other sequences such as human keratins and common lab contaminants. Sequest HT searches were performed using a 20 ppm precursor ion tolerance and requiring each peptide’s N-/C-termini to adhere with trypsin protease specificity, while allowing up to two missed cleavages. Cysteine carbamidomethyl (+57.021) was set as a static modification while methionine oxidation (+15.99492 Da) was set as a variable modification. A MS2 spectra assignment false discovery rate (FDR) of 1% on protein level was achieved by applying the target-decoy database search. Filtering was performed using a Percolator (64bit version). For quantification, a 0.02 m/z window centered on the theoretical m/z value of each the six reporter ions and the intensity of the signal closest to the theoretical m/z value was recorded. Reporter ion intensities were exported in result file of Proteome Discoverer 2.1 search engine as an excel tables.

3.4.20 Co-culturing of E. faecalis and E. lenta A2

E. faecalis MMH594 WT or tyrDC mutant or E. lenta A2 or E. lenta DSM2243 were grown anaerobically for 48 hours from single colonies in BHI medium at 37 °C. These starter cultures were normalized to an OD$_{600}=0.1$ by dilution into fresh BHI medium. 10 µL of each normalized starter culture was diluted in 5 mL of BHI medium containing 0.75% (w/v) arginine and either 1 mM $d_3$-phenyl-L-dopa or dopamine. Cultures were grown for 48 hours at 37 °C and harvested by
centrifugation. Culture supernatants were diluted 1:10 in LC-MS grade methanol and were centrifuged to pellet precipitates. Supernatants were analyzed by LC-MS using Method C described above.

3.4.21 Gain-of-function assays in complex human gut microbiota samples

*E. faecalis* MMH594 WT, *E. faecalis* MMH594 tyrDC mutant, or *E. lenta* A2 were grown anaerobically for 48 hours from single colonies in BHI medium at 37 °C. These starter cultures were normalized to an OD$_{600}$=0.1 by dilution into fresh BHI medium. 10 µL of each normalized starter culture was diluted in MEGA medium containing 1 mM $d_3$-phenyl-L-dopa. At the time of addition of *E. faecalis* and/or *E. lenta*, fecal samples determined to be non-metabolizers with regard to L-dopa were inoculated into the medium as described above and were grown anaerobically for 72 hours at 37 °C. Cultures were harvested by centrifugation, and culture supernatants were diluted 1:10 in LC-MS grade methanol, followed by another round of centrifugation to pellet precipitates. Supernatants were analyzed by LC-MS using Method C described above.

3.4.22 qPCR assays on DNA from bacterial communities

gDNA was extracted from the culture pellets generated in ex vivo experiments using the DNeasy UltraClean Microbial Kit. 2 ng of the extracted DNA from each culture was used for qPCR assays containing 10 µL of iTaq Universal SYBRgreen Supermix (Bio-rad, catalog 3: 1725121), 7 µL of water, and 10 µM each of forward and reverse primers. PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 95 °C for 5 minutes 34 cycles of 95 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min. The program ended with a final extension at 34 °C for 5 mins. The primers used were 16S rRNA primers for *E. lenta* (56): 5’ CAGCAGGGAAGAAATTCGAC 3’ and 5’ TTGAGCCCTCGGATTAGAGA 3’; primers for
dopamine dehydroxylase: 5’ GAGATCTGGTCCACCGTCAT 3’ and 5’ AGTGGAAAGTACACCGGGATG 3’.

3.4.23 Ex vivo assays for dopamine dehydroxylation by LC-MS and Sanger sequencing

Fecal slurries from human donors were prepared as described in the ‘enrichment culturing’ section above. These slurries were diluted 1:100 into BHI medium containing 1% arginine (w/v) and 10 mM formate as well as 500 µM dopamine. Cultures were grown anaerobically at 37 °C for 72 hours. Cultures were harvested by centrifugation, and culture supernatants were diluted 1:10 in LC-MS grade methanol, followed by another round of centrifugation to pellet precipitates. Supernatants were analyzed by LC-MS using Method A described above. gDNA was extracted from the culture pellet using the DNeasy UltraClean Microbial Kit. 1 ng of the extracted DNA from each culture was used for PCR assays containing 10 µL of Phusion High-Fidelity PCR Master mix with HF buffer (NEB, catalog# M0531L), 7 µL of water, and 10 µM each of forward and reverse primers. The primers used to amplify the full-length dopamine dehydroxylase from these samples were 5’ ATGGGTAACCTGACCATG 3’ and 5’ TTACTCCCTCCCTTCGTA 3’. PCR was performed on a C1000 Touch Thermocycler (Bio-Rad), using the following program: initial denaturation at 98 ºC for 30 s, 34 cycles of 98 ºC for 10 s, 61 ºC for 15 s, 72 ºC for 2.5 mins. The program ended with a final extension at 72 ºC for 5 mins. Amplicons were purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE healthcare, catalog# 28-9034-70) and were sequenced using Sanger sequencing (Eton Biosciences) for the region containing the SNP at position 506 using primers 5’ GGGGTGTCCATGTTGCCG GT 3’ and 5’ ACCGGCTACGGCAACGGC 3’. Sequence chromatograms were analyzed in Ape Plasmid Editor (version 2.0.47), and the single nucleotide polymorphism (SNP) at position 506 was called by visual inspection compared to results obtained from control cultures of E. lenta strains. Samples
where two peaks existed at position 506 were determined to have a mix of SNPs present in the sample and were removed from analysis.

**3.4.24 Metagenomic analysis of *E. lenta/dadh* and *E. faecalis/tyrDC* abundance and prevalence across human patients**

A curated collection of human gut microbiomes representing 1870 individuals (116), was used to correlate the abundances of *E. lenta/dadh* and *Enterococcus/tyrDC* (Pearson correlation, R). Prevalence was estimated as a function of rolling minimum abundance cut off. SNP analysis was carried out as before (37), by mapping reads from a set of 96 samples with high *E. lenta* genome coverage to the reference genome of A2 (Assemble accession GCA_003340125.1) after quality filtering with FastP (117) and using Bowtie 2.3.4.1 (112) and SAMtools 1.9 (118). This analysis was done by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).

**3.4.25 Evaluation of inhibitors towards growth and dopamine metabolism in *E. lenta***

*E. lenta* was grown for 48 hours in BHI medium anaerobically at 37 °C to generate turbid starter cultures. The starter culture was diluted 1:10 in 180 μL of fresh BHI medium containing 500 μM L-dopa or dopamine and varying concentrations of carbidopa or AFMT. Cultures were grown anaerobically at 37 °C for 48 hours and harvested by centrifugation. Supernatants were diluted 1:10 in LC-MS grade methanol and were analyzed by LC-MS using Method C described above. Growth was monitored using a Synergy HTX Multi-Mode Microplate Reader (BioTek) by measuring absorbance at 600 nm. Experiments were performed anaerobically, and cultures were grown in 96-well plates (VWR, catalog# 29442-054).
3.4.26 Gnotobiotic mouse experiment with *E. faecalis* and *E. lenta*

3.4.26.1 Colonization of germ-free mice with *E. faecalis* and *E. lenta*

This experiment was performed by Elizabeth Bess (Turnbaugh lab, UCSF). In an anaerobic chamber, wells of a 96-well plate were filled with 100 µL of BHI supplemented with L-cysteine-HCl (0.05%, w/v), L-arginine (1%, w/v), menadione (1 µg/mL), and hemin (5 µg/mL) (this medium was named BHI++) were inoculated with a single colony from agar plates of *Eggerthella lenta* A2, *Eggerthella lenta* DSM 2243, *Enterococcus faecalis* MMH594 WT, *Enterococcus faecalis* MMH594 ΔtyrDC. Cultures were grown for ~24 hr to achieve dense cultures. Then, these cultures were used to inoculate 5 mL of sterile BHI++ medium (anaerobic and in anaerobic chamber) by adding 50 µL of dense culture. Cultures were incubated for approximately 24 hr at 37 °C. Germ-free male Balb/c mice were gavaged with 200 µL of the *E. fecalis* WT (10^7 total inoculation), *E. faecalis* mutant (10^7 total inoculation), *E. lenta* A2 (10^7 total inoculation), *E. lenta* DSM2243 (10^8 total inoculation). There were four colonization groups: *E. faecalis* MMH594 WT and *E. lenta* A2, *E. faecalis* tyrDC mutant and *E. lenta* A2, *E. faecalis* tyrDC mutant and *E. lenta* DSM2243, and *E. faecalis* MMH594 WT and *E. lenta* DSM2243. Mice were allowed to feed *ad libitum* a standard chow. All groups were housed in separate isolators. On day 5, mice were weighed, and fecal pellets were collected. Mice were then orally gavaged with a combined bolus of 10 mg of L-DOPA (phenyl-d₃) and 30 mg of carbidopa, suspended in a solution of 0.25% carboxymethylcellulose sodium salt (filtered sterilized prior to adding drugs), per kg of body weight. Mice were then singly housed in metabolic cages (Hatteras Instruments) with unrestricted access to water and standard chow. Mice were maintained in the metabolic cages for 18 h during which time urine and fecal pellets were separately collected. Mice were euthanized via CO₂ at ~19.5 hrs following oral gavage. Serum from cardiac puncture was collected. Ileal (I8), cecal, and
colon (proximal) tissue and the contents, thereof, and brain were harvested and flash-frozen in liquid nitrogen. The collected urine volume was recorded and then the urine was subjected to the colorimetric assay for catechol detection.

3.4.26.2 Assessment of colonization levels of gnotobiotic mice using qPCR

gDNA was extracted from gnotobiotic mouse stool samples (collected prior to dosing with L-dopa and carbidopa) using the DNeasy Powersoil Kit (Qiagen, catalog# 12888-100). 2 ng of the extracted DNA from each culture was used for qPCR assays containing 10 μL of iTaq Universal SYBRgreen Supermix (Bio-rad, catalog 3: 1725121), 7 μL of water, and 10 μM each of forward and reverse primers. PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 95 °C for 5 minutes 34 cycles of 95 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min. The program ended with a final extension at 34 °C for 5 min. The primers used were 16S rRNA primers for E. lenta (45): The primers used were: 16S primers for E. faecalis (119): 5’ CGCTTCTTTTCTCCGGAGT 3’ and 5’ GCCATGCGGCATAAACTG 3’; primers for tyrDC (120): 5’ CGTACACATTAGTTGCATGGCAT 3’ and 5’ ATGTCTACTTCTCTCCTCCCATTG 3’.

3.4.26.3 Extraction of enzyme-hydrolyzed urine

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. Urine was first spun down at max speed for 10 minutes at 4 °C. 80 μL of urine supernatant was then mixed with 72 μL sodium acetate butter (0.2 M, pH 5.5, containing 1 μM each of non-deuterated standards of L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA) as well as with 10 μL of H. pomatia glucuronidase preparation (Sigma-Aldrich, catalog# G7017-2mL) and 10 μL of H. pomatia sulfatase preparation (Sigma-Aldrich, catalog# S975-5mL). This enzyme-buffer-urine mix was incubated for 18 hours at 37 °C.
to release the free metabolites. To quench this reaction mixture, 200 µL of LC-MS grade MeOH containing 0.1% (w/v) ascorbic acid and 25 µM EDTA, and 300 µL of chloroform were added. Samples were vortexed briefly and then incubated at –80 ºC for 10 minutes to precipitate proteins. Samples were then centrifuged to separate top aqueous layer and bottom chloroform layer, for 10 minutes at 4 ºC. The top aqueous layer was transferred (200 µL for all samples) to a new Eppendorf tube. To this, 250 µL of LC-MS grade MeOH was added followed by incubation at –80 ºC for 50 min, followed by centrifugation. The supernatant was transferred to a different Eppendorf tube. This supernatant was then subjected to evaporation on a Genevac (EZ-2 Elite personal evaporator), HPLC setting (1hr organic 1hr water, no heating). The concentrate was dissolved in 20 µL of MQ water containing 0.1% formic acid, spun down, and 10 µL was analyzed on LC-MS using Method B described above.

3.4.26.4 Extraction of acid-hydrolyzed urine

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. Urine was first spun down at max speed for 10 minutes at 4ºC. 80 µL of supernatant was then mixed with 100 µL 0.5M HCl containing 1 µM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA. This mixture was boiled at 95 ºC for 10 minutes, followed by centrifugation for 10 minutes. Then, 10 µL of sodium acetate butter (0.2M, pH 5.5) was added to stabilize the pH. To this, 800 µL of MeOH containing 0.1% (w/v) ascorbic acid and 25 µM of EDTA were added. The mixture was left at –20ºC for 20 minutes followed by centrifugation to pellet precipitates. This supernatant was then subjected to evaporation on the Genevac (EZ-2 Elite personal evaporator), HPLC setting (1hr organic 1hr water, no heating). The concentrate was dissolved in 200 µL of MQ water containing
0.1% formic acid, spun down, and 10 µL of the resulting supernatant was analyzed on LC-MS using Method B.

3.4.26.5 Extraction of enzyme-hydrolyzed cecal contents

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. The wet weight of cecal contents was first recorded. The cecal contents were then subjected to lyophilization and the dry weight was then recorded. To the lyophilized material, 200 µL of sodium acetate butter (0.2 M, pH 5.5, containing 1 µM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA) was added, as well as 72 of µL H. pomatia glucuronidase preparation (Sigma-Aldrich, catalog# G7017-2mL) and 72 µL H. pomatia sulfatase preparation (Sigma-Aldrich, catalog# S975-5mL). Samples were vortex to homogenize contents. This cecal content-enzyme-buffer mix was incubated for 16 hours at 37ºC to release the free metabolites, followed by centrifugation to pellet particulates. To quench the reaction, to 300 µL of supernatant, of 500 µL LC-MS grade MeOH containing 0.1% (w/v) ascorbic acid and 25 µM EDTA was added to precipitate proteins. Following centrifugation, 500 µL of supernatant was mixed with 500 µL of chloroform. The mixture was left at -20 ºC for 2 hours to allow layers to separate. Samples were then centrifuged to further facilitate separation of layers. 500 µL of the top aqueous layer was transferred to an Eppendorf tube and was then subjected to evaporation on a Genevac (EZ-2 Elite personal evaporator), HPLC setting (1 h organic 1 h water, no heating). The concentrate was dissolved in 600 µL of MQ water containing 0.1% formic acid, spun down, and 10 µL of the supernatant was analyzed on LC-MS using Method B described above.
3.4.26.6 Extraction of enzyme-hydrolyzed stool

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. The wet weight of stool samples was first recorded. The stool samples were then subjected to lyophilization and the dry weight was then recorded. To the lyophilized material, 500 µL of sodium acetate butter (0.2 M, pH 5.5, containing 1 µM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA) was added, as well as 72 µL of *H. pomatia* glucuronidase preparation (Sigma-Aldrich, catalog# G7017-2mL) and 72 µL of *H. pomatia* sulfatase preparation (Sigma-Aldrich, catalog# S975-5mL). Samples were vortex to homogenize contents. This stool-enzyme-buffer mix was incubated for 16 hours at 37ºC to release the free metabolites, followed by centrifugation to pellet particulates. To 300 µL of supernatant, of 500 µL of LC-MS grade MeOH containing 0.1% (w/v) ascorbic acid and 25 µM EDTA was added to precipitate proteins. Following centrifugation, 500 µL of supernatant was mixed with 500 µL of chloroform. The mixture was left at –20 ºC for 2 hours to allow layers to separate. Samples were then centrifuged to further facilitate separation of layers. 500 µL of the top aqueous layer was transferred to an Eppendorf tube and was then subjected to evaporation on a Genevac (EZ-2 Elite personal evaporator), HPLC setting (1 h organic 1 h water, no heating). The concentrate was dissolved in 600 µL of MQ water containing 0.1% formic acid, spun down, and 10 µL of the supernatant was analyzed on LC-MS using Method B described above.

3.4.26.7 Extraction of unhydrolyzed stool

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. The wet weight of stool samples was first recorded. The stool samples were then subjected to lyophilization and the dry weight was then recorded. To the lyophilized material, 500 µL of sodium acetate butter (0.2 M, pH 5.5, containing 35 µM EDTA,
0.1% (w/v) ascorbic acid, 1 µM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA) was added. Samples were vortex to homogenize contents, followed by centrifugation to pellet particulates. To 400 µL of supernatant, of 500 µL LC-MS grade of MeOH containing 0.1% (w/v) ascorbic acid and 25 µM of EDTA was added to precipitate proteins. 500 µL of chloroform was then added. The mixture was left at –80 °C for 10 minutes to allow layers to separate. Samples were then centrifuged to further facilitate separation of layers. The top aqueous layer was transferred to an Eppendorf tube. To this, 500 µL of LC-MS grade MeOH was added, followed by centrifugation. The entire supernatant was subjected to evaporation on a Genevac (EZ-2 Elite personal evaporator), HPLC setting (1.45 h organic 1 h water, no heating). The concentrate was dissolved in 200 µL of MQ water containing 0.1% formic acid, spun down, and 10 µL of the supernatant was analyzed on LC-MS using Method B described above.

3.4.26.8 Extraction of cardiac puncture serum

All manipulations were performed in 1.5 mL Eppendorf tubes. Procedures were performed on ice to prevent oxidation of catechols. Cardiac puncture serum was first spun down to access the plasma. 30 µL of plasma was added to a mixture of 125 µL of 0.2 M sodium acetate buffer (0.2 M, pH 5.5), 250 µL of LC-MS grade MeOH (containing 35 µM EDTA, 0.1% (w/v) ascorbic acid, 1 µM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA), and 500 µL of chloroform. The mixture was vortexed until the solution appeared homogenous and samples were then left at –80 °C for 10 minutes. Organic and aqueous layers were then separated using centrifugation for 10 minutes, and the top aqueous layer was then carefully removed (approximately 50 µL) and transferred to a new Eppendorf tube. This aqueous layer was evaporated using a Genevac (EZ-2 Elite personal evaporator, HPLC setting with 30 mins each for the first and second phases, no heating). The dried residue was then resuspended in 12.5
of μL MilliQ water (containing 0.1 % formic acid) by vortexing. Samples were spun down to pellet potential particulates and 10 μL of the resulting supernatant was injected for LC-MS analysis using Method B described above.

3.4.26.9 Extraction of unhydrolyzed ileal contents

Procedures were performed on ice to prevent oxidation of catechols. All manipulations were performed in 1.5 mL Eppendorf tubes. The wet weight of ileal contents was first recorded. The ileal contents were then subjected to lyophilization and the dry weight was then recorded. To the lyophilized material, 300 μL of sodium acetate butter (0.2 M, pH 5.5, containing 35 μM EDTA, 0.1% (w/v) of ascorbic acid, 1 μM of the non-deuterated standards L-dopa, dopamine, m-tyramine, DOPAC, m-HPAA, and HVA) was added. Samples were vortex to homogenize contents, followed by centrifugation to pellet particulates. To 400 μL of supernatant, of 500 μL of LC-MS grade MeOH containing was added to precipitate proteins. Samples were then centrifuged to pellet precipitates. The top aqueous layer was transferred to an Eppendorf tube. To this, 500 μL of LC-MS grade MeOH was added, followed by centrifugation. The entire supernatant was subjected to evaporation on a Genevac (EZ-2 Elite personal evaporator), HPLC setting (1.45 h organic 1 h water, no heating). The concentrate was dissolved in 200 μL of MQ water containing 0.1% formic acid, spun down, and 10 μL of the supernatant was analyzed on LC-MS using Method B described above.
3.5 References


Chapter 4. A widely distributed metalloenzyme class enables gut microbial metabolism of
host-and diet-derived catechols

This chapter was adapted in part from:

V. Maini Rekdal, P. Nol Bernadino, M. U. Luescher, S. Kiamehr, C. Le, J. E. Bisanz, P. J.
Turnbaugh, E. N. Bess, E. P. Balskus, A widely distributed metalloenzyme class enables gut
4.1 Introduction

Dopamine is not the only catechol metabolized by the human gut microbiota. The catechol structural motif (1,2-dihydroxylated aromatic rings), is found in a diverse range of compounds, including host hormones, dietary phytochemicals, clinically used drugs, and microbial siderophores (1-4) (Figure 4.1). Like dopamine, many of these catecholic molecules undergo dehydroxylation by the gut microbiota, a reaction that selectively replaces the para hydroxyl group of the catechol with a hydrogen atom (5) (Figure 4.2). This uniquely microbial reaction was discovered over six decades ago, when studies revealed that rats and rabbits fed catechols excreted dehydroxylated metabolites in their urine (6). The reaction is notable not only because of its prominence in gut microbial metabolism, but also because it is particularly challenging due to the stability of the aromatic ring system (5, 6).

![Figure 4.1 Classes of catechol compounds encountered by the human gut microbiota. The catechol motif is displayed in red.](image-url)
Prominent substrates for gut microbial dehydroxylation include fostamatinib (7), a spleen tyrosine kinase inhibitor approved in 2018 for the treatment of immune thrombocytopenia, the neurotransmitters dopamine and norepinephrine (8, 9), as well as the phytochemicals ellagic acid (found in nuts and berries) (10, 11), caffeic acid (a universal lignin precursor in plants) (12-14), hydrocaffeic acid (derived from microbial metabolism of xenobiotics) (14, 15), lignans (from flaxseed) (16, 17), and catechin (present in chocolate and tea) (Figure 4.2) (18). Some of these compounds are directly dehydroxylated, while others undergo host or microbial reactions prior to dehydroxylation (Figure 4.2). For example, dehydroxylation is the last reaction in the multi-step, interspecies metabolism of dietary lignans into the phytoestrogen enterodiol (19). Similarly, caffeic acid is first reduced into hydrocaffeic acid prior to dehydroxylation (15). Hydrocaffeic acid is also a central intermediate in gut microbial pathways that break down the drug L-dopa (as described in Chapter 3) and diet-derived aromatic compounds such as polyphenols (20-24) (Figure 4.2). These known examples of catechol dehydroxylation were discovered on a case-by-case basis rather than through systematic screening efforts. The prominence of catechols in diverse compounds encountered by gut bacteria suggests that the scope of microbial catechol dehydroxylation may include molecules beyond those already known.
Figure 4.2 Examples of catechol dehydroxylation reactions performed by the human gut microbiota.
Dehydroxylation alters the bioactivity of catechol compound and produces metabolites that act both locally in the gut and systemically to influence human health and disease (Figure 4.2). For example, gut microbes transform ellagic acid into urolithins, a group of metabolites found at micromolar levels in the serum that may account for the beneficial effects of pomegranate, nuts, and berries (25). Urolithin A induces mitophagy and promotes lifespan in the model organism *C. elegans* and increases muscle function in rodents (26). This molecule also activates aryl hydrocarbon receptor-nuclear factor erythroid 2-related factor 2-dependent pathways, reducing inflammation and promoting gut barrier function in mice (27). Enterodiol, a dehydroxylated metabolite derived from flaxseed lignans, may have similarly beneficial effects. Microbial production of this phytoestrogen is associated with decreased tumor burden in mice and lowered breast cancer risk in post-menopausal women, and the dehydroxylated metabolites directly inhibit growth of anthracene-induced mammary carcinomas in rats (28-33). Similarly, dehydroxylated metabolites of diet-derived catechins increase the activity of CD4+ T cells and enhance the cytotoxic activities of NK cells in vivo, potentially explaining the links between catechin consumption and lowered cancer risk (34, 35). Finally, *m*-hydroxyphenylpropionic acid (*m*-hppa), the product of hydrocaffeic acid dehydroxylation, decreases arterial blood pressure and increases bone mass in rats (36, 37). While the biological effects of many dehydroxylated metabolites remain poorly understood, the excretion of these molecules in urine suggest potential interactions with host processes throughout the body (7, 9, 38). By changing levels of the parent compound and
producing metabolites with altered activities, microbial catechol dehydroxylation has implications for host-microbe interactions, nutrition, and pharmacology.

Previous work has revealed gut bacterial strains that dehydroxylate catechols (Figure 4.2). For example, various strains of the prevalent gut microbe *Eggerthella lenta* and its close relatives metabolize catechin through benzyl ether reduction followed by catechol dehydroxylation (18), and we uncovered the role of these organisms in the dehydroxylation of dopamine (described in Chapter 3 of this thesis). Similarly, *Gordonibacter* and *Ellagibacter*, two organisms closely related to *E. lenta*, dehydroxylate dietary ellagic acid to produce urolithins, and *Gordonibacter* also dehydroxylates dietary lignans in the production of enterodiol (19, 39, 40). All the identified catechol dehydroxylating human gut bacterial strains are members of the Actinobacterial phylum, suggesting this group of organisms may be broadly relevant to this activity in the human gut (Figure 4.2). Finally, while a *Pseudomonas* sp. isolated from rat feces was reported in 1966 to dehydroxylate caffeic acid into *m*-hydroxyphenylpropionic acid, *m*-coumaric acid, and other unidentified products, no human gut organism has been demonstrated to perform this reaction (13). This latter example highlights that many catechol dehydroxylation reactions remain understood only at the level of the microbial community and have not yet been assigned to an individual isolate in the human gut microbiota. More importantly, at the time we began our work, the genetic and enzymatic basis of catechol dehydroxylation in the gut microbiota was unknown. This represented a significant gap in our understanding of microbial metabolism and has made it challenging to mechanistically investigate potential links between microbial catechol dehydroxylation and host and microbial biology.

As reported in Chapter 3, we discovered a catechol dehydroxylating enzyme from the prevalent human gut Actinobacterium *Eggerthella lenta*. This enzyme participates in an
interspecies gut microbial pathway that degrades the Parkinson’s disease medication L-dopa, catalyzing the regioselective $p$-dehydroxylation of dopamine to yield $m$-tyramine (29). To identify the dopamine dehydroxylase enzyme, we grew $E.\ lenta$ strain A2 with and without dopamine and used RNA sequencing (RNA-seq) to identify genes differentially regulated by dopamine. Only 15 genes were significantly upregulated in the presence of dopamine, including a putative molybdenum-dependent enzyme that was induced $>2500$ fold. Hypothesizing that this gene encoded the dopamine dehydroxylase, we purified the enzyme from $E.\ lenta$ and confirmed its activity in vitro. The dopamine dehydroxylase (Dadh) identified through these experiments is predicted to bind bis-molybdopterin guanine dinucleotide (bis-MGD), a complex metallocofactor that coordinates a catalytically essential molybdenum atom (41). Our earlier work also demonstrated that a Single Nucleotide Polymorphism (SNP) in Dadh correlates with metabolism by isolated strains and complex microbial communities. Here, in this chapter, we sought to explore the substrate scope of Dadh and its broader role in catechol dehydroxylation by the gut microbiota.
4.2 Results

4.2.1 Dadh specifically metabolizes catecholamines that are available in the gut

Because the human gut microbiota metabolizes a range of catecholic compounds (Figure 4.2), we first investigated whether the recently discovered Dadh possessed promiscuous dehydroxylase activity. We evaluated the reactivity of natively purified *E. lenta* A2 Dadh towards a panel of established or potential host- and diet-derived catechol substrates (Table 4.1 and Figure 4.3). This enzyme displayed a narrow substrate scope, metabolizing only dopamine and the structurally related neurotransmitter norepinephrine, which differ only by the presence of a benzylic hydroxyl group (Figure 4.4). To identify the structural elements necessary for substrate recognition by Dadh, we profiled its activity towards synthetic and commercially available dopamine analogs (1-14) (Table 4.2 and Figure 4.5). We acknowledge Dr. Michael Luescher in our laboratory for making and characterizing the synthetic dopamine analogs. We found that Dadh tolerated only minor modifications to the dopamine scaffold, including a single *N*-methylation and the presence of additional hydroxyl groups on the aromatic ring (Figure 4.5). The catechol moiety was absolutely necessary for activity, and dehydroxylation required that at least one hydroxyl group to be in the *para* position relative to the amine substituent. These data demonstrated that Dadh specifically recognizes the catecholamine scaffold and that as-yet-unidentified enzymes likely metabolize other catechols in the human gut microbiota.
Table 4.1 Physiologically relevant catechol substrates used in assays with natively purified dopamine dehydroxylase from *E. lenta* A2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Natural origin</th>
<th>Catalog number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caffeic acid</td>
<td><img src="image1" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# C0625-2G</td>
</tr>
<tr>
<td>Hydrocaffeic acid</td>
<td><img src="image2" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# 102601-10G</td>
</tr>
<tr>
<td>(-)-Epicatechin</td>
<td><img src="image3" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# E1753-1G</td>
</tr>
<tr>
<td>(+)-Catechin</td>
<td><img src="image4" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# C1251-5G</td>
</tr>
<tr>
<td>Protocatechuic acid</td>
<td><img src="image5" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# 37580-25G-F</td>
</tr>
<tr>
<td>DOPAC</td>
<td><img src="image6" alt="Structure" /></td>
<td>Host</td>
<td>Millipore Sigma, catalog# 850217-1G</td>
</tr>
<tr>
<td>Ellagic acid</td>
<td><img src="image7" alt="Structure" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# E4642-5G</td>
</tr>
<tr>
<td>D-L-epinephrine</td>
<td><img src="image8" alt="Structure" /></td>
<td>Host</td>
<td>Millipore Sigma, catalog# E4642-5G</td>
</tr>
<tr>
<td>D-L-norepinephrine</td>
<td><img src="image9" alt="Structure" /></td>
<td>Host</td>
<td>Millipore Sigma, catalog# A7256-1G</td>
</tr>
<tr>
<td>Dopamine</td>
<td><img src="image10" alt="Structure" /></td>
<td>Host</td>
<td>Millipore Sigma, catalog# PHR1090-1G</td>
</tr>
</tbody>
</table>
Figure 4.3 SDS-PAGE of natively purified dopamine dehydroxylase from *E. lenta* A2. Ladder is the Precision Plus Protein™ All Blue Standards (first lane from the left), while the subsequent lane represents the combined dopamine-dehydroxylating fractions from the size exclusion column, the last chromatography step of the activity-based purification from *E. lenta* A2. The dopamine dehydroxylase (115 kDa predicted size) is highlighted with the red asterisk.
Figure 4.4 Dadh specifically metabolizes catecholamines and does not dehydroxylate other physiologically relevant catechol substrates.

Enzyme (0.1 µM) was incubated with substrate (500 µM) for 22 hours at room temperature, followed by analysis using LC-MS. Bars represent the mean ± the standard error (SEM) of three biological replicates (enzyme reactions).
Table 4.2 Commercially available and synthesized dopamine analogs used in assays with natively purified dopamine dehydroxylase from *E. lenta A2.*

<table>
<thead>
<tr>
<th>Analog#</th>
<th>Name</th>
<th>Structure</th>
<th>Catalog number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>p</em>-Tyramine</td>
<td><img src="image1.png" alt="Structure of p-Tyramine" /></td>
<td>Sigma Aldrich, catalog# T2879-1G</td>
</tr>
<tr>
<td>2</td>
<td>3-Aminotyramine</td>
<td><img src="image2.png" alt="Structure of 3-Aminotyramine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>3</td>
<td>3-Methoxytyramine</td>
<td><img src="image3.png" alt="Structure of 3-Methoxytyramine" /></td>
<td>Sigma Aldrich, catalog# M4251-100MG</td>
</tr>
<tr>
<td>4</td>
<td>2,3-Dihydroxyphenylethylamine</td>
<td><img src="image4.png" alt="Structure of 2,3-Dihydroxyphenylethylamine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>5</td>
<td>4,6-Dihydroxyphenylethylamine</td>
<td><img src="image5.png" alt="Structure of 4,6-Dihydroxyphenylethylamine" /></td>
<td>Enamine, catalog # EN300-65185</td>
</tr>
<tr>
<td>6</td>
<td>1,4-Dihydroxyphenylethylamine</td>
<td><img src="image6.png" alt="Structure of 1,4-Dihydroxyphenylethylamine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>7</td>
<td>3,5-Dihydroxyphenylethylamine</td>
<td><img src="image7.png" alt="Structure of 3,5-Dihydroxyphenylethylamine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>8</td>
<td>Hydroxytyrosol</td>
<td><img src="image8.png" alt="Structure of Hydroxytyrosol" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>9</td>
<td>3,4-Dihydroxybenzylamine</td>
<td><img src="image9.png" alt="Structure of 3,4-Dihydroxybenzylamine" /></td>
<td>Sigma Aldrich, catalog# 858781-250MG</td>
</tr>
<tr>
<td>10</td>
<td><em>N</em>-Methyldopamine</td>
<td><img src="image10.png" alt="Structure of N-Methyldopamine" /></td>
<td>Santa Cruz, catalog# sc-358430A</td>
</tr>
<tr>
<td>11</td>
<td>2-Hydroxydopamine</td>
<td><img src="image11.png" alt="Structure of 2-Hydroxydopamine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
<tr>
<td>12</td>
<td>6-Hydroxydopamine</td>
<td><img src="image12.png" alt="Structure of 6-Hydroxydopamine" /></td>
<td>Sigma Aldrich, catalog #H4381-100MG</td>
</tr>
<tr>
<td>13</td>
<td>5-Hydroxydopamine</td>
<td><img src="image13.png" alt="Structure of 5-Hydroxydopamine" /></td>
<td>Sigma Aldrich, catalog# 151564-100G</td>
</tr>
<tr>
<td>14</td>
<td>2,3,5-Trihydroxyphenylethylamine</td>
<td><img src="image14.png" alt="Structure of 2,3,5-Trihydroxyphenylethylamine" /></td>
<td>Synthesized in our laboratory</td>
</tr>
</tbody>
</table>
Figure 4.5 Dadh tolerates only minor changes to the catecholamine scaffold.  
A) Dopamine analogs evaluated in this study. B) Reactivity towards dopamine analogs. Enzyme (0.1 µM) was incubated with substrate (500 µM) for 22 hours at room temperature, followed by analysis using LC-MS. Bars represent the mean ± the standard error (SEM) of three biological replicates (enzyme reactions).
The narrow substrate scope of Dadh and its apparent specificity for catecholamines prompted us to explore whether the transcriptional regulation of Dadh displayed similar specificity. As described in Chapter 3 of this thesis, we had previously observed that dopamine upregulates dadh in E. lenta. To test the specificity of this transcriptional response, we cultured E. lenta A2 in the presence of a subset of the dopamine analogs that we had tested with the purified enzyme, measured dehydroxylation using LC-MS, and profiled the global transcriptome using RNA-seq. We found that the regulation of dadh was also specific for the catecholamine scaffold (Figure 4.6). While the catecholamines dopamine and norepinephrine induced dadh expression (>2500-fold and >700-fold, respectively) and were dehydroxylated by E. lenta, analogs lacking the catechol (analog 1) or having a shorter side chain (analog 9) did not elicit a transcriptional or metabolic response (Figure 4.6 and Table 4.3) (see Chapter 3 for RNA-sequencing data in response to dopamine). Together with our biochemical results, these transcriptional data suggest that Dadh may have evolved for the purpose of catecholamine neurotransmitter metabolism in E. lenta. We propose that dopamine is an endogenous substrate of this enzyme because it is the best substrate for the Dadh enzyme both in vitro and in vivo, induces the highest levels of expression in E. lenta, and is produced at substantial levels within the human gastrointestinal tract (42).
Figure 4.6 The catecholamine scaffold is required for dadh induction and dehydroxylation by E. lenta A2 cells.
A) Structures of analogs evaluated in transcriptional and whole cell assays. B) Specificity in metabolism and dadh transcription. Transcriptional induction was assessed using RNA-seq, with the fold induction shown on the x-axis (foldchange>2, FDR<0.01). To assess whole-cell metabolism, E. lenta was grown anaerobically for 48 hours in BHI medium with 500 µM of each substrate, and the culture supernatant was analyzed for dehydroxylated metabolites using a colorimetric assay or LC-MS. RNA-sequencing data represent the log2fold change from n=3 independent cultures for each condition (compound/vehicle). The metabolism data represent the mean ± the SEM of three biological replicates (independent bacterial cultures).
Table 4.3 Genes differentially expressed upon exposure of *E. lenta* A2 to norepinephrine.
The data represent the genes that are differentially expressed in response to exposure of *E. lenta* A2 to 0.5 mM DL-norepinephrine relative to vehicle (>2-fold difference, FDR<0.1). The dopamine dehydroxylase (*dadh*) is highlighted in red. Data are from n=3 bacterial cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_00671</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>11.41</td>
<td>1.08E-25</td>
</tr>
<tr>
<td>Elenta-A2_00535</td>
<td>molydopterin_dinucleotide-binding_region (<em>dadh</em>)</td>
<td>9.6</td>
<td>6.64E-13</td>
</tr>
<tr>
<td>Elenta-A2_00533</td>
<td>hypothetical_protein</td>
<td>7.14</td>
<td>5.34E-13</td>
</tr>
<tr>
<td>Elenta-A2_00999</td>
<td>hypothetical_protein</td>
<td>2.2</td>
<td>2.37E-13</td>
</tr>
<tr>
<td>Elenta-A2_01000</td>
<td>hypothetical_protein</td>
<td>2.19</td>
<td>1.36E-15</td>
</tr>
<tr>
<td>Elenta-A2_00997</td>
<td>DBCCT_transporter</td>
<td>2.15</td>
<td>3.73E-10</td>
</tr>
<tr>
<td>Elenta-A2_00998</td>
<td>Creatininase</td>
<td>2.01</td>
<td>5.47E-06</td>
</tr>
<tr>
<td>Elenta-A2_00996</td>
<td>General_substrate_transporter</td>
<td>1.61</td>
<td>8.79E-07</td>
</tr>
<tr>
<td>Elenta-A2_01462</td>
<td>hypothetical_protein</td>
<td>1.43</td>
<td>3.08E-09</td>
</tr>
<tr>
<td>Elenta-A2_01463</td>
<td>4Fe-4S_ferredoxin_iron-sulfur_binding_domain_protein</td>
<td>1.43</td>
<td>4.98E-07</td>
</tr>
<tr>
<td>Elenta-A2_01840</td>
<td>Integrase_core_domain_protein</td>
<td>1.42</td>
<td>1.04E-07</td>
</tr>
<tr>
<td>Elenta-A2_01464</td>
<td>molydopterin_dinucleotide-binding_region</td>
<td>1.35</td>
<td>6.36E-16</td>
</tr>
<tr>
<td>Elenta-A2_02451</td>
<td>DDEAD/DEAH_box_helicase_domain_protein</td>
<td>1.25</td>
<td>1.18E-25</td>
</tr>
<tr>
<td>Elenta-A2_01910</td>
<td>molydopterin_dinucleotide-binding_region</td>
<td>-1.12</td>
<td>1.90E-19</td>
</tr>
<tr>
<td>Elenta-A2_01559</td>
<td>membrane_bound_O-acyl transferase</td>
<td>-1.17</td>
<td>2.43E-07</td>
</tr>
<tr>
<td>Elenta-A2_00170</td>
<td>flavocytochrome_c</td>
<td>-1.79</td>
<td>3.16E-10</td>
</tr>
</tbody>
</table>
4.2.2 Dopamine promotes gut bacterial growth by serving as an alternative electron acceptor

The specificity of Dadh for dopamine suggested this metabolic activity might have an important physiological role in *E. lenta*. As described already in Chapter 3, we noted the chemical parallels between catechol dehydroxylation and reductive dehalogenation, a metabolic process in which halogenated aromatics serve as alternative electron acceptors in certain environmental bacteria (43). We next tested the hypothesis that dopamine dehydroxylation could provide a similar growth benefit to gut bacteria. We observed no growth increase when *E. lenta* was grown in complex BHI medium containing dopamine relative to a vehicle control (water) (Figure 4.7). In contrast, we found that including dopamine in a minimal medium lacking electron acceptors (basal medium, the medium used during enrichment culturing to identify a dopamine dehydroxylating strain) increased the endpoint optical density of *E. lenta* cultures (Figure 4.8). The endpoint OD\textsubscript{600} was significantly lower in the basal medium compared to BHI, indicating a general lack of nutrients for robust growth. However, his growth-promoting effect was only observed in dopamine-metabolizing *E. lenta* strains; non-metabolizing strains that express an apparently inactive enzyme did not gain a growth advantage (Figure 4.8). The effect of dopamine on *E. lenta* growth contrasts with recent studies of digoxin, a drug that is reduced by *E. lenta* without impacting growth in the same basal medium examined here (44).
Figure 4.7 Dopamine does not promote growth of *E. lenta* A2 in BHI medium. *E. lenta* A2 was grown anaerobically in BHI medium at 37 °C with and without dopamine. The data shown are the mean ± the SEM (n=3 replicate growth experiments). The experiment was performed once.

Figure 4.8 Dopamine specifically promotes growth of dopamine-dehydroxylating *E. lenta* strains in basal medium. Strains were grown anaerobically in basal medium containing 10 mM sodium acetate for 48-72 hours at 37 °C before growth and metabolism were assessed. Bars represent the mean ± the SEM of three biological replicates (bacterial cultures). A) shows endpoint growth and B) shows endpoint metabolism, as assessed by the catechol colorimetric assay.
We further investigated the relationship between dopamine and bacterial growth in the metabolizing strain *E. lenta* A2. The growth increase observed in response to dopamine was dose-dependent (Figure 4.9), supporting the idea that electron acceptors are limiting in this medium. The growth increase also mirrored the effects of the known electron acceptors DMSO and nitrate (44, 45) and did not derive from the product of dopamine dehydroxylation, *m*-tyramine (Figure 4.10). Additionally, the growth benefit was directly tied to dopamine dehydroxylation. Inclusion of tungstate in the growth medium, which inactivates the big-MGD cofactor of Dadh, blocked metabolism and inhibited the growth increase. In contrast, inclusion of molybdate in the growth medium did not impact growth or metabolism (Figure 4.10). Molybdate and tungstate alone did not impact *E. lenta* A2 growth in the basal medium (Figure 4.11). Taken together, these results indicate that active metabolism of dopamine provides a growth advantage to *E. lenta*, likely by serving as an alternative electron acceptor.

![Figure 4.9 Dopamine promotes growth of *E. lenta* A2 in basal medium in a dose-dependent manner.](image)

*E. lenta* was grown anaerobically in basal medium containing 10 mM sodium acetate at 37 °C with varying concentrations of dopamine for 48 hours. Bars represent the mean ± the SEM (n=3 replicate cultures for each strain).
Tungstate inhibits dopamine-dependent growth and dopamine metabolism by *E. lenta* A2 grown in basal medium.

*E. lenta* was grown anaerobically for 48 hours at 37 °C in basal medium containing 10 mM sodium acetate. Dopamine, *m*-tyramine, and sodium nitrate were added to a final concentration of 1 mM, while DMSO was added to a final concentration of 14 mM at the time of inoculation. Sodium tungstate (WO$_4^{2-}$) and sodium molybdate (MoO$_4^{2-}$) were added to a final concentration of 0.5 mM. Bars represent the mean ± the SEM of three biological replicates (bacterial cultures). The experiment was performed twice. A) shows the endpoint growth and B) shows the endpoint metabolism in the presence of dopamine, as assessed by the colorimetric assay for catechol detection.
Figure 4.11 Tungstate and molybdate alone do not impact *E. lenta* A2 growth in basal medium.

*E. lenta* was grown anaerobically in basal medium containing 10 mM sodium acetate at 37 °C with either 1 mM dopamine, sodium tungstate, or sodium molybdate (0.5 mM each) for 36 hours before growth was assessed. Bars represent the mean ± the SEM of three biological replicates (independent cultures).

We next examined whether dopamine could promote *E. lenta* growth in microbial communities. First, we competed dopamine metabolizing and non-metabolizing *E. lenta* strains in minimal medium. *E. lenta* is genetically intractable, preventing the use of engineered plasmids encoding defined fluorescence or antibiotic resistance as markers of strain identity. Instead, we took advantage of intrinsic differences in tetracycline (Tet) resistance and pigment production to differentiate the closely related strains in pairwise competitions (46). Across competition experiments using two different strain combinations, inclusion of dopamine in growth medium significantly increased the proportion of the metabolizer relative to the non-metabolizer (Figures 4.12 and 4.13). This was driven by the growth increase of the metabolizer rather than a decrease
in growth of the non-metabolizer, indicating that the growth benefit of the metabolizer rather than toxicity towards the non-metabolizer explains the altered strain ratio (Figures 4.12 and 4.13).

Figure 4.12 The dopamine metabolizer *E. lenta* A2 outcompetes the non-metabolizer *E. lenta* DSM2243 in the presence of dopamine in basal medium. Strains were grown together for 72 hours at 37 °C in basal medium containing 10 mM acetate and were then plated. Pigment production was used to determine strain identity, with A2 forming white colonies, and DSM2243 forming pink colonies. A) Raw CFU counts of each strain in the competition experiment. B) Proportion of the metabolizing A2 strain. Bars represent the mean ± the SEM of six biological replicates.
Figure 4.13 The dopamine metabolizer *E. lenta* Valencia outcompetes the non-metabolizer *E. lenta* W1BHI6 in the presence of dopamine in basal medium. 

*E. lenta* strains were in basal medium containing 10 mM acetate. Strains were grown together for 72 hours at 37 °C and were then plated on BHI medium. Antibiotic resistance was used to determine strain identity. Bars represent the mean ± the SEM of six biological replicates (bacterial cultures).
Next, we explored the impact of dopamine on Tet-resistant *E. lenta* in the presence of a defined bacterial community representing the major phylogenetic diversity in the human gut (Table 4.4 and Figure 4.14) (47, 48). Including dopamine in the medium boosted the growth of metabolizers in this model community by an order of magnitude while non-metabolizing strains did not gain a growth advantage (Figure 4.14). Finally, we evaluated the impact of dopamine on *E. lenta* strains present in complex human gut microbiotas. We cultured fecal samples from 24 unrelated subjects *ex vivo* in the presence and absence of dopamine and used qPCR to assess the abundance *E. lenta* and *dadh*. Both *dadh* and *E. lenta* significantly increased by an order of magnitude in cultures containing dopamine (p<0.005, two-tailed unpaired t-test) (Figure 4.15). Finally, we amplified the full length *dadh* gene from these cultures and sequenced the region harboring the SNP that distinguishes metabolizing and non-metabolizing strains. These assays indicated that the increase in *dadh* abundance in the complex communities was accompanied by a shift from a mixture of inactive and active *dadh* variants to a dominance of the metabolizing R506 variant (p<0.01, Fisher’s exact test) (Figure 4.16). Finally, we noticed in these growth assays that a small number of samples did not display an increase in *E. lenta* or *dadh* abundance (n=4 and n=3 samples, respectively) (Figure 4.15). While the factors influencing this outcome are unclear, they could include the possibility that these specific communities support the growth of *E. lenta* in other ways so that dopamine metabolism does not provide any additional advantage, that these samples contain inhibitory factors, or that organisms not targeted by our primers were responsible for metabolism. Altogether, these results are consistent with the hypothesis that dopamine dehydroxylation can increase the fitness of metabolizing *E. lenta* strains in microbial communities.
Table 4.4 Bacterial strains used in this study.

<table>
<thead>
<tr>
<th>Bacterial strains used in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggerthella lenta 1-60</td>
</tr>
<tr>
<td>Eggerthella lenta 11C</td>
</tr>
<tr>
<td>Eggerthella lenta 1-3-56</td>
</tr>
<tr>
<td>Eggerthella lenta 14A</td>
</tr>
<tr>
<td>Eggerthella lenta 22C</td>
</tr>
<tr>
<td>Eggerthella lenta 28B</td>
</tr>
<tr>
<td>Eggerthella lenta 32-6-1 6 NA</td>
</tr>
<tr>
<td>Eggerthella lenta AB12 #2</td>
</tr>
<tr>
<td>Eggerthella lenta AB8 #2</td>
</tr>
<tr>
<td>Eggerthella lenta AN51LG</td>
</tr>
<tr>
<td>Eggerthella lenta CC7/5 D5 2</td>
</tr>
<tr>
<td>Eggerthella lenta CC8/2 BHI2</td>
</tr>
<tr>
<td>Eggerthella lenta CC8/6 D5 4</td>
</tr>
<tr>
<td>Eggerthella lenta DSM 11767</td>
</tr>
<tr>
<td>Eggerthella lenta DSM 11863</td>
</tr>
<tr>
<td>Eggerthella lenta DSM 15644</td>
</tr>
<tr>
<td>Eggerthella lenta DSM2243</td>
</tr>
<tr>
<td>Eggerthella lenta MR #12</td>
</tr>
<tr>
<td>Eggerthella lenta RC4/6F</td>
</tr>
<tr>
<td>Eggerthella lenta Valencia</td>
</tr>
<tr>
<td>Eggerthella lenta W1 BHI 6</td>
</tr>
<tr>
<td>Eggerthella sinesis DSM 16107</td>
</tr>
<tr>
<td>Eggerthella lenta A2</td>
</tr>
<tr>
<td>Gordonibacter pamelaeae 28C</td>
</tr>
<tr>
<td>Gordonibacter pamelaeae 3C</td>
</tr>
<tr>
<td>Enterococcus faecalis OG1RF</td>
</tr>
<tr>
<td>Eschericia coli MG1655</td>
</tr>
<tr>
<td>Bacteroides fragilis ATCC 25285</td>
</tr>
<tr>
<td>Clostridium sporogenes ATCC 15579</td>
</tr>
<tr>
<td>Edwarsiella tarda ATCC 23685</td>
</tr>
<tr>
<td>Paraeggerthella hongkongensis RC2/2A</td>
</tr>
</tbody>
</table>
Figure 4.14 Dopamine promotes growth of metabolizing *E. lenta* strains in basal medium in the presence of a defined microbial community.

Tetracycline resistant (TetR) *E. lenta* strains were grown with tetracycline sensitive (TetS) gut isolates for 48 hours at 37 °C in basal medium containing 10 mM acetate. Plating on BHI medium containing tetracycline allowed enumeration of *E. lenta*. Bars represent the mean ± the SEM of three biological replicates (bacterial cultures).
Figure 4.15 Dopamine increases *dadh* and *E. lenta* abundance in complex human gut microbiotas cultured in basal medium ex vivo.

Samples from unrelated individuals (n = 24) were grown for 72 hours at 37 °C in basal medium containing 10 mM sodium acetate with or without dopamine and qPCR was used to assess abundance of *dadh* and *E. lenta*. Two individuals were excluded from this analysis as they did not demonstrate quantitative metabolism of dopamine after incubation. Each point represents a different individual. Lines connect data from the same individual between the two conditions. (***P*=0.0005, two-tailed unpaired t-test, n=22 samples per group. (**P*<0.005, two-tailed unpaired t-test, n=22 per group). A) represents *dadh* abundance, while B) represents *E. lenta* abundance.
Figure 4.16 Dopamine increases the proportion of metabolizing dadh variants in complex human gut microbiotas cultured in basal medium ex vivo.

The same gDNA used in Figure 4.15 was used to amplify full-length dadh and determine the SNP status at position 506 using Sanger sequencing. As in Figure 4.15, two individuals were removed prior to analysis as they did not demonstrate quantitative metabolism of dopamine after incubation. In addition, one individual was not included in this analysis due to failure of obtaining high quality sequencing data. (**P=0.008, Fisher’s exact test, n=9 CGC samples and n=12 CGC/AGC samples for vehicle; n=18 CGC samples and n=3 CGC/AGC samples for dopamine).
4.2.3 A screen of human gut Actinobacteria uncovers dehydroxylation of host and plant-derived catechols

Having uncovered Dadh’s specialized role in gut bacterial dopamine metabolism, we sought to identify gut bacterial strains and enzymes that could dehydroxylate other catechol substrates. Among human gut bacteria, only *Eggerthella* and closely related members of the Actinobacteria phylum have been reported to perform catechol dehydroxylation. For example, *Eggerthella* metabolizes dopamine (as reported in Chapter 3) and (+)-catechin (18), while related *Gordonibacter* species dehydroxylate ellagic acid (39) and didemethylsecoisolariciresinol (dmSECO), an intermediate in the multi-step biosynthesis of the anti-cancer metabolite enterodiol (19). These reports suggest that Actinobacteria could be a promising starting point to identify new dehydroxylating strains and enzymes. Thus, we screened a library of related gut Actinobacteria (46) (n=3 replicates for each strain) for metabolism of a range of compounds relevant in the human gut, including plant- and host-derived small molecules, bacterial siderophores, and FDA-approved catecholic drugs (1-3) (Table 4.5 and Figure 4.17). We initially used a colorimetric assay that detects catechols to assess metabolism, which allowed us to rapidly screen for potential catechol depletion across the collection of 25 strains. For ellagic acid, we used LC-MS to detect metabolism. The strain incubations and colorimetric assays were performed by Paola Nol (laboratory of Elizabeth Bess, UC Irvine).

We observed catechol dehydroxylation by most strains that we tested. However, this activity was limited to plant- and host-derived molecules. Drugs and siderophores were not metabolized by our strain library, although we cannot rule out the possibility that these compounds were metabolized at levels below the limit of detection of the colorimetric assay or are transformed by complex gut microbial communities. Overall, catechol depletion differed greatly between
strains, with relatively minor differences in substrate structure having large consequences for apparent activity (Figure 4.17). For instance, *Gordonibacter pamelaeae* 3C completely depleted the catechol 3,4-dihydroxyphenylacetic acid (DOPAC) but did not display any notable activity towards dihydromandelic acid, which has an additional benzylic hydroxyl group. Similarly, although hydroxytyrosol and dopamine share the same carbon scaffold, differing only by the presence of a primary alcohol vs. amine, only dopamine was clearly depleted by at least one strain in this initial screen (Figure 4.17).
Table 4.5 Catechol substrates used in screen of gut Actinobacteria for catechol metabolism.  
3,4-DHMA = 3,4-dihydroxymandelic acid; 2,3-DHBA = 2,3-dihydroxybenzoic acid.

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Natural origin</th>
<th>Catalog number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocatechuic acid</td>
<td><img src="image1" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# 37580-25G-F</td>
</tr>
<tr>
<td>Hydrocaffeic acid</td>
<td><img src="image2" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# 102601-10G</td>
</tr>
<tr>
<td>Caffeic acid</td>
<td><img src="image3" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# C0625-2G</td>
</tr>
<tr>
<td>(+)-Catechin</td>
<td><img src="image4" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millimore Sigma, catalog# C1251-5G</td>
</tr>
<tr>
<td>(+/-)-Catechin</td>
<td><img src="image5" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# C1788-500MG</td>
</tr>
<tr>
<td>(-)-Epicatechin</td>
<td><img src="image6" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# E1753-1G</td>
</tr>
<tr>
<td>Ellagic acid</td>
<td><img src="image7" alt="Structural formula" /></td>
<td>Plant</td>
<td>Millipore Sigma, catalog# E2250-1G</td>
</tr>
<tr>
<td>Hydroxytyrosol</td>
<td><img src="image8" alt="Structural formula" /></td>
<td>Plant</td>
<td>Ava Chem Scientific, catalog# 2528</td>
</tr>
<tr>
<td>DOPAC</td>
<td><img src="image9" alt="Structural formula" /></td>
<td>Host</td>
<td>Millipore Sigma, catalog# 850217-1G</td>
</tr>
<tr>
<td>3,4-DHMA</td>
<td><img src="image10" alt="Structural formula" /></td>
<td>Host</td>
<td>Carbo Synth, catalog# FD22118</td>
</tr>
<tr>
<td>Dopamine</td>
<td><img src="image11" alt="Structural formula" /></td>
<td>Host</td>
<td>Sigma-Aldrich, catalog# H8502-25G</td>
</tr>
<tr>
<td>L-norepinephrine</td>
<td><img src="image12" alt="Structural formula" /></td>
<td>Host</td>
<td>Matrix Scientific, catalog# 037592-500MG</td>
</tr>
<tr>
<td>L-epinephrine</td>
<td><img src="image13" alt="Structural formula" /></td>
<td>Host</td>
<td>Alfa Aesar, catalog# L04911.06</td>
</tr>
<tr>
<td>L-dopa</td>
<td><img src="image14" alt="Structural formula" /></td>
<td>FDA-approved drug</td>
<td>Oakwood Chemical, catalog# 358380-25g</td>
</tr>
<tr>
<td>Methyldopa</td>
<td><img src="image15" alt="Structural formula" /></td>
<td>FDA-approved drug</td>
<td>Chemcruz, catalog# sc-203092</td>
</tr>
<tr>
<td>Carbidopa</td>
<td><img src="image16" alt="Structural formula" /></td>
<td>FDA-approved drug</td>
<td>Sigma-Aldrich, catalog# PHR1655-1G</td>
</tr>
<tr>
<td></td>
<td>Chemical Structure</td>
<td>Source and Catalog Number</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Fenoldopam</strong></td>
<td><img src="image" alt="Fenoldopam" /></td>
<td>FDA-approved drug Sigma Aldrich, catalog# SML0198-10MG</td>
<td></td>
</tr>
<tr>
<td><strong>Apomorphine</strong></td>
<td><img src="image" alt="Apomorphine" /></td>
<td>FDA-approved drug Sigma Aldrich, catalog# A4393-100MG</td>
<td></td>
</tr>
<tr>
<td><strong>Enterobactin</strong></td>
<td><img src="image" alt="Enterobactin" /></td>
<td>Microbial siderophore Gift from professor Liz Nolan, MIT</td>
<td></td>
</tr>
<tr>
<td><strong>2,3-DHBA</strong></td>
<td><img src="image" alt="2,3-DHBA" /></td>
<td>Microbial siderophore Millipore Sigma, catalog# 126209-5G</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.17 Colorimetric screen for catechol dehydroxylation by human gut Actinobacteria.
A) Examples of catechol compounds used in our screen. B) Metabolism of catechols. For all compounds except ellagic acid, individual strains were grown in the presence of a single catechol substrate for 24-48 hours at 37 °C in BHI medium containing 10 mM formate and 1% arginine (w/v). Metabolism was assessed using the catechol colorimetric assay. For ellagic acid, strains were grown in for 72 hours at 37 °C in BHI medium. Metabolism was assessed using LC-MS. Data from bacterial incubations were normalized to the sterile control. Data represent the mean of three biological replicates (independent bacterial cultures). A white box means that the strain was not screened for metabolism of the specific compound. The screen was performed by Paola Nol (laboratory of Elizabeth Bess, UC Irvine).
Our initial screen used a colorimetric method for catechol detection. However, apparent catechol depletion could result from catechol dehydroxylation, import of the compound into the cell, or degradation by other means. Thus, to follow up on our colorimetric screen, we first spot-checked select candidate metabolizing strains and used LC-MS to confirm direct production of metabolites. While some substrates were dehydroxylated directly, others underwent stepwise metabolism (Figures 4.18 and 4.19). For instance, (+)-catechin metabolism by *E. lenta* A2 involved benzyl ether reduction and dehydroxylation to give bicyclic derivative 2, as has been observed previously (18) (Figure 4.18). Similarly, ellagic acid metabolism in *G. pamelaeae* 3C was also carried out by multi-step pathway involving hydrolysis and decarboxylation followed by multiple dehydroxylations to give Urolithin C (39) (Figure 4.19). Some *Eggerthella* strains reduced the alkene of caffeic acid before dehydroxylation, while other strains dehydroxylated this diet-derived metabolite directly to produce *m*-coumaric acid (Figure 4.18). Both of these transformations were previously observed in rat cecal contents incubated with caffeic acid ex vivo (12). The host-derived compounds DOPAC and norepinephrine were directly dehydroxylated by *Gordonibacter* and *Eggerthella*, respectively (Figures 4.18 and 4.19). While (+)-catechin and ellagic acid metabolism was previously linked to human gut Actinobacteria, the dehydroxylation of the neurotransmitter norepinephrine, plant-derived hydrocaffeic acid, and the dopamine metabolite DOPAC were only previously observed in complex communities. Culture supernatants from potential metabolizers incubated with 3,4-dihydroxybenzoic acid, 2,3-dihydroxybenzoic acid, L-dopa, carbidopa, and epinephrine were also analyzed by LC-MS, but direct dehydroxylation was never detected.
Figure 4.18 Confirmation of catechol dehydroxylation by *E. lenta* strains.

A) Pathway for (+)-catechin metabolism by *E. lenta* strains. The pathway proceeds by initial benzyl ether reduction, followed by dehydroxylation of the bicyclic catechol intermediate.

---

B) Pathway for (+)-catechin metabolism by *E. lenta* strains. The pathway proceeds by initial benzyl ether reduction, followed by dehydroxylation of the bicyclic catechol intermediate.

---

C) Graph showing the mass peak area of m/z 275.0925 for (+)-catechin and its metabolites.

D) Graph showing the mass peak area of m-HPPA (m-Hydroxyphenylpropionic acid) for caffeic acid and its metabolites.

E) Graph showing the mass peak area of m-HPPA (m-Hydroxyphenylpropionic acid) for caffeic acid and its metabolites.

F) Graph showing the mass peak area of m-HPPA (m-Hydroxyphenylpropionic acid) for hydrocaffeic acid and its metabolites.

G) Graph showing the mass peak area of m-Octopamine for noradrenaline and its metabolites.

---

Figure 4.18 Confirmation of catechol dehydroxylation by *E. lenta* strains.

A) Pathway for (+)-catechin metabolism by *E. lenta* strains. The pathway proceeds by initial benzyl ether reduction, followed by dehydroxylation of the bicyclic catechol intermediate.
Figure 4.18 (Continued).

B) Pathway for caffeic acid and hydrocaffeic acid metabolism by *E. lenta* strains. Caffeic acid can either be directly dehydroxylated into *m*-coumaric acid or can be reduced to hydrocaffeic acid (12, 13). Hydrocaffeic acid is directly dehydroxylated into *m*-hydroxyphenylpropionic acid (*m*-hppa).

C) Metabolism of (+)-catechin by *E. lenta* A2. This strain was grown in BHI medium with and without 0.5 mM (+)-catechin for 48 hours at 37 °C. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM mass peak area of the Extracted Ion Chromatogram (EIC) for the bicyclic, dehydroxylated derivative 2 shown in A) (three biological replicates).

D) Direct dehydroxylation of caffeic acid by *E. lenta* 28b. This strain was grown in BHI medium with and without 0.5 mM caffeic acid for 48 hours at 37 °C. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM mass peak area of *m*-coumaric acid (three biological replicates).

E) Metabolism of caffeic acid by *E. lenta* RC4/6F. This strain was grown in BHI medium with and without 0.5 mM caffeic acid for 48 hours at 37 °C. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM mass peak area of *m*-hppa (three biological replicates). This strain reduced caffeic acid into hydrocaffeic acid, which was then directly dehydroxylated.

F) Metabolism of hydrocaffeic acid by *E. lenta* A2. This strain was grown in BHI medium with and without 0.5 mM hydrocaffeic acid for 48 hours at 37 °C. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM concentration of *m*-hppa (three biological replicates).

G) Metabolism of DL-norepinephrine by *E. lenta* 28b. This strain was grown in BHI medium with and without 0.5 mM DL-norepinephrine for 48 hours at 37 °C. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM mass peak area of *m*-octopamine (three biological replicates).
**Figure 4.19 Confirmation of catechol dehydroxylation by* G. pamelaeae.**

A) Pathway for ellagic acid metabolism by *Gordonibacter* strains. The pathway proceeds by initial hydrolysis and decarboxylation, followed by two subsequent dehydroxylation reactions. B-D) Metabolism of ellagic acid by *G. pamelaeae* 3C. This strain was grown in BHI medium with and without 0.5 mM ellagic acid for 48 hours at 37 ºC.
Figure 4.19 (Continued).

Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM mass peak area of the EIC for the structures shown in A) (three biological replicates). The majority of ellagic acid was converted fully to Urolithin C. E) Metabolism of DOPAC by \textit{G. pamelaecae} 3C. This strain was grown in BHI medium with and without 1 mM DOPAC for 48 hours at 37 ºC. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM concentration of the metabolite \textit{m}-hydroxyphenylacetic acid (\textit{m}-HPAA) resulting from direct DOPAC dehydroxylation (three biological replicates, e.g. bacterial cultures).

Having verified in a small subset of strains that potential catechol dehydroxylation reactions identified in our colorimetric screen indeed do take place, we decided to further analyze and characterize the dehydroxylation of hydrocaffeic acid, (+)-catechin, and DOPAC. Instead of culturing only a subset of bacteria, we repeated the incubations of our entire strain library with these compounds and used LC-MS/MS to confirm the production of dehydroxylated metabolites. Consistent with our initial colorimetric screen, this metabolite analysis showed a large variability in these activities across closely related gut bacterial strains (Figure 4.20). For example, only \textit{Gordonibacter} and \textit{Paraeggerthella} strains (n=3 total) dehydroxylated DOPAC, while hydrocaffeic acid was dehydroxylated by almost all strains across the library, including \textit{Eggerthella}, \textit{Paraeggerthella}, and \textit{Gordonibacter}. (+)-catechin metabolism was confined to \textit{E. lenta}. However, no strain dehydroxylated all three compounds. This variability suggests that distinct enzymes might dehydroxylate different catechols and reinforces the idea that microbial phylogeny is not predictive of metabolic function (44, 49).
Gut Actinobacteria display strain-level variability in the dehydroxylation of DOPAC, hydrocaffeic acid, and (+)-catechin.

To evaluate catechol metabolism, individual strains were grown in triplicate in the presence of a single catechol substrate for 48 hours at 37 °C in BHI medium. Metabolism was assessed using LC-MS/MS. Bars represent the mean ± the SEM of three biological replicates.
4.2.4 Impact of DOPAC, catechin, and hydrocaffeic acid on gut Actinobacterial growth in basal medium

We next asked whether the new catechol dehydroxylations identified in our screen could promote growth of metabolizing gut bacteria in a manner similar to dopamine. Growing *E. lenta* A2 with 1 mM hydrocaffeic acid or (+)-catechin in the basal medium lacking other electron acceptors, we found that these catechol compounds provided a growth advantage similar to dopamine (Figure 4.21). This suggests that catechols more broadly can act as alternative electron acceptors for *Eggerthella*.

However, when we grew the DOPAC-metabolizing strain *G. pamelaeae* 3C with 1 mM DOPAC under the same conditions, we found no clear effect on endpoint optical density, likely because metabolism did not proceed substantially under these conditions (Figure 4.22). Substituting 10 mM acetate for 10 mM formate restored metabolism in this medium and revealed a more striking difference in endpoint optical density between DOPAC and the vehicle control (Figure 4.22). Nonetheless, the maximum growth was even lower than that observed with *E. lenta* and there was no obvious growth in the vehicle control, making us question whether the results reflected a real biological difference. Repeating the experiment, we turned to spot plating and colony counts as an alternative method to measure growth, and our results pointed in the direction of a potential growth advantage (Figure 4.22). However, we believe that further work is necessary to optimize the medium for robust *G. pamelaeae* growth and to establish the extent to which DOPAC dehydroxylation provides this organism with a significant growth advantage.
**Figure 4.21** Hydrocaffeic acid and (+)-catechin promotes growth of *E. lenta* A2 in basal medium.

A) Growth of *E. lenta* A2 with and without 1 mM hydrocaffeic acid in basal medium containing 10 mM acetate. The strain was grown anaerobically for 48 hours at 37 °C before growth was assessed. Bars represent the mean ± the SEM of three biological replicates. B) Growth of *E. lenta* A2 with and without 1 mM (+)-catechin in basal medium containing 10 mM acetate. The strain was grown anaerobically for 48 hours at 37 °C before growth was assessed. Bars represent the mean ± the SEM of three biological replicates.
Figure 4.22 DOPAC may promote growth of *G. pamelaeae* 3C in the presence of formate in basal medium.

A) Growth and B) metabolism of *G. pamelaeae* 3C grown with and without 1 mM DOPAC in basal medium containing 10 mM acetate. The strain was grown anaerobically for 72 hours at 37 °C before growth was assessed using a spectrophotometer. Metabolism was assessed using the colorimetric method for catechol detection. Bars represent the mean ± the SEM of three biological replicates. C) Growth and D) metabolism of *G. pamelaeae* 3C grown with and without 1 mM DOPAC in basal medium containing 10 mM formate. The strain was grown anaerobically for 72 hours at 37 °C before growth was assessed using a spectrophotometer. Metabolism was assessed using the colorimetric method for catechol detection. Bars represent the mean ± the SEM of three biological replicates. D) Growth of *G. pamelaeae* 3C in basal medium with 10 mM formate, as assessed by spot plating method. *G. pamelaeae* 3C grown with and without 1 mM DOPAC in basal medium containing 10 mM acetate. The strain was grown anaerobically for 72 hours at 37 °C before growth was assessed using the spot plating method. Bars represent the mean ± the SEM of three biological replicates
4.2.5 Catechol dehydroxylation by gut Actinobacteria is inducible and specific

We next sought to determine the molecular basis of the dehydroxylation reactions examined above (DOPAC, hydrocaffeic acid, and (+)-catechin). To test the hypothesis that specific rather than promiscuous enzymes were involved, we first established that dehydroxylation is an inducible activity in *Gordonibacter* and *Eggerthella* strains (Figure 4.23). This allowed us to use the dehydroxylase activity of cell lysates as a proxy for transcriptional induction and a means of examining dehydroxylase activity. We grew *E. lenta* A2 in the presence of (+)-catechin, hydrocaffeic acid, and dopamine, and grew *G. pamelaeae* 3C in the presence of DOPAC. We then screened each anaerobic lysate for its activity towards all four substrates individually. Consistent with our prediction, each lysate quantitively dehydroxylated only the catechol substrate with which the strain had been grown. While the *E. lenta* lysates did not display any promiscuity, cell lysate from *G. pamelaeae* grown in the presence of DOPAC displayed reduced activity (<45% conversion) towards hydrocaffeic acid, which structurally resembles DOPAC (Figure 4.24). Overall, these results suggest that different catechol substrates induce the expression of distinct dehydroxylase enzymes that are specific in their activity and transcriptional regulation.
Figure 4.23 Catechol dehydroxylation by whole cell suspensions of *E. lenta* A2 and *G. pamelaee* 3C is inducible and oxygen sensitive.
Figure 4.23 (Continued).

A) Inducibility and oxygen sensitivity of dopamine dehydroxylase activity in *E. lenta* A2. *E. lenta* A2 was grown anaerobically in BHI medium containing 1% arginine and 10 mM formate. 0.5 mM of dopamine was added to induce dehydroxylase expression, followed by pelleting of cells and resuspension in PBS. Whole cell suspensions were exposed to dopamine (500 μM) or vehicle (water), and reactions were allowed to proceed anaerobically or aerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Bar graph displays the mass peak area for a single sample.

B) Inducibility and oxygen sensitivity of hydrocaffeic acid dehydroxylase activity in *E. lenta* A2. *E. lenta* A2 was grown anaerobically in BHI medium containing 1% arginine and 10 mM formate. 0.5 mM of hydrocaffeic acid was added to induce dehydroxylase expression, followed by pelleting of cells and resuspension in PBS. Whole cell suspensions were exposed to hydrocaffeic acid (500 μM) or vehicle (water) and reactions were allowed to proceed anaerobically or aerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Bar graph displays the mass peak area of *m*-hydroxyphenylpropionic acid (*m*-HPPA) for a single sample.

C) Inducibility and oxygen sensitivity of catechin dehydroxylase activity in *E. lenta* A2. *E. lenta* A2 was grown anaerobically in BHI medium containing 1% arginine and 10 mM formate. 0.5 mM of (+)-catechin was added to induce dehydroxylase expression, followed by pelleting of cells and resuspension in PBS. Whole cell suspensions were exposed to (+)-catechin (500 μM) or vehicle (DMF) and reactions were allowed to proceed anaerobically or aerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Bar graph displays the mass peak area of the dehydroxylated catechin derivative in a single sample.

D) Inducibility and oxygen sensitivity of (+)-catechin benzyl ether reduction activity in *E. lenta* A2. *E. lenta* A2 was grown anaerobically in BHI medium containing 1% arginine and 10 mM formate. 0.5 mM of (+)-catechin was added to induce dehydroxylase expression, followed by pelleting of cells and resuspension in PBS. Whole cell suspensions were exposed to (+)-catechin (500 μM) or vehicle (DMF) and reactions were allowed to proceed anaerobically or aerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Bar graph displays the mass peak area of the benzyl ether reduced catechin derivative in a single sample.

Benzyl ether reduction was constitutive but oxygen sensitive in *E. lenta* A2.

E) Inducibility and oxygen sensitivity of DOPAC acid dehydroxylase activity *G. pamelaeeae* 3C. *G. pamelaeeae* 3C was grown anaerobically in BHI medium containing 10 mM formate. 0.5 mM of DOPAC was added to induce dehydroxylase expression, followed by pelleting of cells and resuspension in PBS. Whole cell suspensions were exposed to DOPAC (500 μM) or vehicle (water) and reactions were allowed to proceed anaerobically or aerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Bar graph displays the mass peak area *m*-hydroxyphenylacetic acid (*m*-HPAA) for a single sample. All experiments described in this figure were performed once.
Figure 4.24 Specificity and inducibility of gut Actinobacterial dehydroxylation as assessed by activity of cell lysates.

A) Specificity of dehydroxylase regulation and activity in *E. lenta* A2. *E. lenta* A2 was grown anaerobically in BHI medium containing 1% arginine and 10 mM formate. 0.5 mM of catechol was added to induce dehydroxylase expression, followed by anaerobic lysis and enzyme assays. Crude lysates were exposed to different substrates (500 µM) and reactions were allowed to proceed anaerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Heat map represents the mean of three biological replicates (lysate reactions). The experiment was performed twice. B) Specificity of DOPAC dehydroxylase regulation and activity in *G. pamelaeae* A2. *G. pamelaeae* 3C was grown anaerobically in BHI medium containing 10 mM formate. 0.5 mM of catechol was added to induce dehydroxylase expression, followed by anaerobic lysis and enzyme assays. Crude lysates were exposed to different substrates (500 µM) and reactions were allowed to proceed anaerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Heat map represents the mean of three biological replicates (lysate reactions). The experiment was performed twice.
4.2.6 Discovery of enzymes that dehydroxylate (+)-catechin and hydrocaffeic acid

We expected that the catechol dehydroxylase enzymes responsible for these additional metabolic activities would likely resemble Dadh, as only molybdenum-dependent enzymes are known to catalyze aromatic dehydroxylation (41, 50, 51). To identify the molecular basis of (+)-catechin and hydrocaffeic acid dehydroxylation in E. lenta A2, we turned to RNA-seq, the approach that we successfully used to identify the dopamine dehydroxylase Dadh. We followed the experimental protocol we used to identify Dadh, growing E. lenta A2 to early exponential phase and adding each catechol substrate, and then harvesting the cells after 1.5 hours of induction (when cells had grown from approximately OD_{600}=0.200 at time of induction to 0.500 at the time of harvest). Hydrocaffeic acid and (+)-catechin each upregulated a number of genes (25 and 41, respectively), including two different predicted molybdenum-dependent enzymes (Tables 4.6 and 4.7). While one of these predicted molybdenum-dependent enzymes was among the highest upregulated genes in response to each substrate (450-fold upregulated in response to catechin, >2000-fold with hydrocaffeic acid), the expression of the other enzyme was only increased 3-fold relative to the vehicle. Thus, we propose that the most highly upregulated molybdenum-dependent enzyme in each dataset is the most reasonable candidate dehydroxylase. The candidate hydrocaffeic acid dehydroxylase (Elenta-A2_02815, named hcdh) shares 35.3% amino acid identity with Dadh, while the candidate (+)-catechin dehydroxylase (E. lenta-A2_00577, named cadh) shares 50.9% amino acid identity with Dadh (Tables 4.6 to 4.8). Similar to Dadh, both Hcdh and Cadh are predicted to be multi-subunit enzymes comprised of a catalytic molybdenum-binding subunit, an electron-shuttling ferredoxin, and a membrane anchor (Tables 4.6 and 4.7).
Table 4.6 Genes differentially expressed upon exposure of *Eggerthella lenta* A2 to hydrocaffeic acid.

The data represent the genes that are differentially expressed in response to 0.5 mM hydrocaffeic acid relative to vehicle control (>2-fold difference, FDR<0.1). The catalytic subunit of the putative hydrocaffeic acid dehydroxylase (*hcdh*), as well as its predicted 4Fe-4S and membrane anchor partners, are highlighted in red. Data are from n=4 bacterial cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_02813</td>
<td>Hypothetical protein (<em>hcdh</em> membrane anchor)</td>
<td>11.21</td>
<td>4.70E-62</td>
</tr>
<tr>
<td>Elenta-A2_02815</td>
<td>Molybdopterin dinucleotide binding region (<em>hcdh</em>)</td>
<td>11.14</td>
<td>4.70E-163</td>
</tr>
<tr>
<td>Elenta-A2_02814</td>
<td>4Fe-4S ferredoxin (<em>hcdh</em> 4Fe4S partner)</td>
<td>10.31</td>
<td>8.00E-53</td>
</tr>
<tr>
<td>Elenta-A2_02810</td>
<td>4Fe-4S ferredoxin</td>
<td>8.31</td>
<td>6.80E-18</td>
</tr>
<tr>
<td>Elenta-A2_02816</td>
<td>Hypothetical protein</td>
<td>7.75</td>
<td>3.10E-12</td>
</tr>
<tr>
<td>Elenta-A2_02811</td>
<td>4Fe-4S ferredoxin</td>
<td>6.83</td>
<td>6.70E-17</td>
</tr>
<tr>
<td>Elenta-A2_02820</td>
<td>Hypothetical protein</td>
<td>6.37</td>
<td>5.30E-47</td>
</tr>
<tr>
<td>Elenta-A2_02819</td>
<td>Cytoplasmic chaperone TorD family protein</td>
<td>6.29</td>
<td>2.70E-52</td>
</tr>
<tr>
<td>Elenta-A2_02817</td>
<td>Hypothetical protein</td>
<td>5.8</td>
<td>5.10E-40</td>
</tr>
<tr>
<td>Elenta-A2_02822</td>
<td>Transcriptional regulator, LuxR family</td>
<td>3.8</td>
<td>5.60E-14</td>
</tr>
<tr>
<td>Elenta-A2_02821</td>
<td>hypothetical_protein</td>
<td>2.7</td>
<td>2.95E-83</td>
</tr>
<tr>
<td>Elenta-A2_02818</td>
<td>Hypothetical protein</td>
<td>2.55</td>
<td>0.00013</td>
</tr>
<tr>
<td>Elenta-A2_01460</td>
<td>formate_dehydrogenase_family_accessory_protein</td>
<td>2.12</td>
<td>1.32E-05</td>
</tr>
<tr>
<td>Elenta-A2_01461</td>
<td>formate_dehydrogenase_accessory_protein</td>
<td>1.67</td>
<td>8.90E-07</td>
</tr>
<tr>
<td>Elenta-A2_01459</td>
<td>peptidase_C60_sortase_A_and_B</td>
<td>1.52</td>
<td>0.004867669</td>
</tr>
<tr>
<td>Elenta-A2_01464</td>
<td>molybdopterin_dinucleotide-binding_region</td>
<td>1.5</td>
<td>3.06E-48</td>
</tr>
<tr>
<td>Elenta-A2_01463</td>
<td>4Fe-4S_ferredoxin</td>
<td>1.44</td>
<td>2.12E-23</td>
</tr>
<tr>
<td>Elenta-A2_01465</td>
<td>molybdopterin_oxidoreductase_Fe4S4_region</td>
<td>1.39</td>
<td>1.76E-14</td>
</tr>
<tr>
<td>Elenta-A2_01462</td>
<td>hypothetical_protein</td>
<td>1.22</td>
<td>1.24E-11</td>
</tr>
<tr>
<td>Elenta-A2_02489</td>
<td>ABC_transporter_related</td>
<td>1.1</td>
<td>7.47E-06</td>
</tr>
<tr>
<td>Elenta-A2_01331</td>
<td>hypothetical_protein</td>
<td>1.08</td>
<td>0.000121478</td>
</tr>
<tr>
<td>Elenta-A2_00815</td>
<td>hypothetical_protein</td>
<td>-1.08</td>
<td>1.49E-05</td>
</tr>
</tbody>
</table>
Table 4.6 (Continued).

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_01910</td>
<td>molydopterin_dinucleotide-binding_region</td>
<td>-1.26</td>
<td>7.51E-17</td>
</tr>
<tr>
<td>Elenta-A2_01911</td>
<td>4Fe-4S_ferredoxin</td>
<td>-1.36</td>
<td>5.47E-16</td>
</tr>
</tbody>
</table>
Table 4.7 Genes differentially expressed genes upon exposure of *Eggerthella lenta* A2 to (+)-catechin.

The data represent the genes that are differentially expressed in response to 0.5 mM (+)-catechin relative to vehicle control (>2-fold difference, FDR<0.1). The catalytic subunit of the putative hydrocaffeic acid dehydroxylase (*cadh*), as well as its predicted 4Fe-4S and membrane anchor partners, are highlighted in red. Data are from n=3 bacterial cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_00233</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>14.38</td>
<td>3.22E-38</td>
</tr>
<tr>
<td>Elenta-A2_00691</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>11.32</td>
<td>2.09E-32</td>
</tr>
<tr>
<td>Elenta-A2_00578</td>
<td>4Fe-4S_ferredoxin_iron-sulfur_binding_domain_protein (cadh 4Fe4S partner)</td>
<td>9.09</td>
<td>1.24E-24</td>
</tr>
<tr>
<td>Elenta-A2_00577</td>
<td>molybdopterin_dinucleotide-binding_region (cadh)</td>
<td>8.85</td>
<td>2.50E-123</td>
</tr>
<tr>
<td>Elenta-A2_00579</td>
<td>hypothetical_protein (cadh membrane anchor)</td>
<td>8</td>
<td>5.40E-60</td>
</tr>
<tr>
<td>Elenta-A2_01215</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>7.55</td>
<td>0</td>
</tr>
<tr>
<td>Elenta-A2_00587</td>
<td>hypothetical_protein</td>
<td>3.99</td>
<td>6.36E-09</td>
</tr>
<tr>
<td>Elenta-A2_01625</td>
<td>NADH:flavin_oxidoreductase/NADH_oxidase</td>
<td>3.99</td>
<td>8.47E-141</td>
</tr>
<tr>
<td>Elenta-A2_01615</td>
<td>Enoyl-CoA_hydratase/isomerase</td>
<td>3.95</td>
<td>5.16E-178</td>
</tr>
<tr>
<td>Elenta-A2_01619</td>
<td>hypothetical_protein</td>
<td>3.91</td>
<td>2.90E-66</td>
</tr>
<tr>
<td>Elenta-A2_01621</td>
<td>Electron_transfer_flavoprotein_alpha/beta_-subunit</td>
<td>3.87</td>
<td>1.60E-20</td>
</tr>
<tr>
<td>Elenta-A2_02817</td>
<td>hypothetical_protein</td>
<td>3.86</td>
<td>5.40E-12</td>
</tr>
<tr>
<td>Elenta-A2_01620</td>
<td>Electron transfer flavoprotein alpha subunit</td>
<td>3.84</td>
<td>5.01E-23</td>
</tr>
<tr>
<td>Elenta-A2_01613</td>
<td>DL-carnitine dehydratase/bile acid inducible protein F</td>
<td>3.82</td>
<td>1.66E-168</td>
</tr>
<tr>
<td>Elenta-A2_01624</td>
<td>acyl-CoA_dehydrogenase_domain_protein</td>
<td>3.79</td>
<td>2.85E-217</td>
</tr>
<tr>
<td>Elenta-A2_01614</td>
<td>NADPH dependent FMN reductase</td>
<td>3.72</td>
<td>6.55E-23</td>
</tr>
<tr>
<td>Elenta-A2_01622</td>
<td>acyl-CoA_dehydrogenase_domain_protein</td>
<td>3.71</td>
<td>7.25E-25</td>
</tr>
<tr>
<td>Elenta-A2_01609</td>
<td>hypothetical_protein</td>
<td>3.22</td>
<td>2.83E-43</td>
</tr>
<tr>
<td>Elenta-A2_01618</td>
<td>hydrolase</td>
<td>3.15</td>
<td>3.25E-60</td>
</tr>
<tr>
<td>Elenta-A2_01608</td>
<td>transcriptional_regulator2C_LysR_family</td>
<td>2.98</td>
<td>3.05E-71</td>
</tr>
<tr>
<td>Elenta-A2_01616</td>
<td>AMP-dependent_synthetase_and_ligase</td>
<td>2.97</td>
<td>1.06E-20</td>
</tr>
<tr>
<td>Elenta-A2_01610</td>
<td>transcriptional_regulator2C_LysR_family</td>
<td>2.8</td>
<td>3.39E-21</td>
</tr>
<tr>
<td>Locus tag</td>
<td>Annotation</td>
<td>log2FoldChange</td>
<td>p-value (FDR&lt;0.1)</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Elenta-A2_01546</td>
<td>thioesterase_superfamily_protein</td>
<td>2.71</td>
<td>8.10E-08</td>
</tr>
<tr>
<td>Elenta-A2_01617</td>
<td>beta-lactamase_domain_protein</td>
<td>2.55</td>
<td>1.09E-04</td>
</tr>
<tr>
<td>Elenta-A2_01493</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>2.54</td>
<td>3.05E-33</td>
</tr>
<tr>
<td>Elenta-A2_01611</td>
<td>major_facilitator_superfamily_MFS_1</td>
<td>2.4</td>
<td>6.45E-60</td>
</tr>
<tr>
<td>Elenta-A2_00576</td>
<td>transcriptional_regulator2C_LuxR_family</td>
<td>2.16</td>
<td>7.78E-19</td>
</tr>
<tr>
<td>Elenta-A2_00581</td>
<td>molybdopterin_dinucleotide-binding_region</td>
<td>1.62</td>
<td>1.75E-10</td>
</tr>
<tr>
<td>Elenta-A2_01602</td>
<td>acyl-CoA_dehydrogenase_domain_protein</td>
<td>1.62</td>
<td>2.17E-21</td>
</tr>
<tr>
<td>Elenta-A2_02346</td>
<td>argininosuccinate_synthase</td>
<td>1.43</td>
<td>2.86E-24</td>
</tr>
<tr>
<td>Elenta-A2_00138</td>
<td>ABC_transporter_related</td>
<td>1.41</td>
<td>8.05E-13</td>
</tr>
<tr>
<td>Elenta-A2_02622</td>
<td>chaperonin_GroEL</td>
<td>1.3</td>
<td>4.58E-60</td>
</tr>
<tr>
<td>Elenta-A2_00137</td>
<td>hypothetical_protein</td>
<td>1.29</td>
<td>1.67E-08</td>
</tr>
<tr>
<td>Elenta-A2_00139</td>
<td>transcriptional_regulator2C_TetR_famil</td>
<td>1.27</td>
<td>1.80E-08</td>
</tr>
<tr>
<td>Elenta-A2_01820</td>
<td>phospho-N-acetylmuramoyl-pentapeptide-transferase</td>
<td>1.23</td>
<td>1.99E-24</td>
</tr>
<tr>
<td>Elenta-A2_01129</td>
<td>MATE_efflux_family_protein</td>
<td>1.21</td>
<td>1.10E-19</td>
</tr>
<tr>
<td>Elenta-A2_02587</td>
<td>hypothetical_protein</td>
<td>1.19</td>
<td>1.09E-34</td>
</tr>
<tr>
<td>Elenta-A2_00167</td>
<td>cell_wall/surface_repeat_protein</td>
<td>1.19</td>
<td>3.32E-13</td>
</tr>
<tr>
<td>Elenta-A2_00256</td>
<td>Rhodanese_domain_protein</td>
<td>1.17</td>
<td>7.25E-22</td>
</tr>
<tr>
<td>Elenta-A2_02345</td>
<td>argininosuccinate_lyase</td>
<td>1.08</td>
<td>1.01E-10</td>
</tr>
<tr>
<td>Elenta-A2_01456</td>
<td>chaperonin_Cpn10</td>
<td>1.08</td>
<td>7.89E-25</td>
</tr>
<tr>
<td>Elenta-A2_01456</td>
<td>aspartate_carbamoyltransferase</td>
<td>1.06</td>
<td>8.19E-10</td>
</tr>
<tr>
<td>Elenta-A2_01326</td>
<td>hypothetical_protein</td>
<td>-1</td>
<td>1.52E-13</td>
</tr>
<tr>
<td>Elenta-A2_02456</td>
<td>hypothetical_protein</td>
<td>-1.01</td>
<td>1.47E-24</td>
</tr>
<tr>
<td>Elenta-A2_00896</td>
<td>hypothetical_protein</td>
<td>-1.02</td>
<td>5.90E-12</td>
</tr>
<tr>
<td>Elenta-A2_02832</td>
<td>hypothetical_protein</td>
<td>-1.02</td>
<td>5.13E-30</td>
</tr>
<tr>
<td>Elenta-A2_01974</td>
<td>hypothetical_protein</td>
<td>-1.12</td>
<td>5.15E-24</td>
</tr>
<tr>
<td>Locus tag</td>
<td>Annotation</td>
<td>log2FoldChange</td>
<td>p-value (FDR&lt;0.1)</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Elenta-A2_02146</td>
<td>hypothetical_protein</td>
<td>-1.12</td>
<td>4.71E-33</td>
</tr>
<tr>
<td>Elenta-A2_00730</td>
<td>polar_amino_acid_ABC_transporter</td>
<td>-1.13</td>
<td>8.82E-09</td>
</tr>
<tr>
<td>Elenta-A2_01591</td>
<td>hypothetical_protein</td>
<td>-1.13</td>
<td>2.10E-13</td>
</tr>
<tr>
<td>Elenta-A2_01359</td>
<td>4Fe-4S_ferredoxin_iron-sulfur_binding_domain_protein</td>
<td>-1.14</td>
<td>1.43E-08</td>
</tr>
<tr>
<td>Elenta-A2_01382</td>
<td>General_secretory_system_II_protein_E_domain_protein</td>
<td>-1.17</td>
<td>1.43E-07</td>
</tr>
<tr>
<td>Elenta-A2_01347</td>
<td>Endoribonuclease_L-PSP</td>
<td>-1.21</td>
<td>1.91E-31</td>
</tr>
<tr>
<td>Elenta-A2_02130</td>
<td>extracellular_solute-binding_protein</td>
<td>-1.21</td>
<td>2.28E-20</td>
</tr>
<tr>
<td>Elenta-A2_02855</td>
<td>chaperone_protein_DnaK</td>
<td>-1.22</td>
<td>8.03E-17</td>
</tr>
<tr>
<td>Elenta-A2_00895</td>
<td>tRNA_pseudouridine_synthase_A</td>
<td>-1.25</td>
<td>1.14E-24</td>
</tr>
<tr>
<td>Elenta-A2_01051</td>
<td>extracellular_solute-binding_protein</td>
<td>-1.26</td>
<td>3.04E-36</td>
</tr>
<tr>
<td>Elenta-A2_00971</td>
<td>ABC_transporter_related</td>
<td>-1.26</td>
<td>5.43E-38</td>
</tr>
<tr>
<td>Elenta-A2_01601</td>
<td>Aldehyde_Dehydrogenase</td>
<td>-1.26</td>
<td>2.44E-15</td>
</tr>
<tr>
<td>Elenta-A2_00527</td>
<td>anaerobic_dimethyl_sulfoxide_reductase</td>
<td>-1.34</td>
<td>1.02E-07</td>
</tr>
<tr>
<td>Elenta-A2_00974</td>
<td>polar_amino_acid_ABC_transporter</td>
<td>-1.39</td>
<td>3.48E-22</td>
</tr>
<tr>
<td>Elenta-A2_02132</td>
<td>hypothetical_protein</td>
<td>-1.43</td>
<td>6.39E-18</td>
</tr>
<tr>
<td>Elenta-A2_02336</td>
<td>hypothetical_protein</td>
<td>-1.44</td>
<td>7.43E-31</td>
</tr>
<tr>
<td>Elenta-A2_01606</td>
<td>membrane_protein_of_unknown_function</td>
<td>-1.44</td>
<td>4.33E-33</td>
</tr>
<tr>
<td>Elenta-A2_02706</td>
<td>two_component_transcriptional_regulator2C</td>
<td>-1.45</td>
<td>5.14E-37</td>
</tr>
<tr>
<td>Elenta-A2_02665</td>
<td>extracellular_solute-binding_protein</td>
<td>-1.46</td>
<td>9.23E-30</td>
</tr>
<tr>
<td>Elenta-A2_01226</td>
<td>ATPase2C_P-type</td>
<td>-1.52</td>
<td>7.75E-18</td>
</tr>
<tr>
<td>Elenta-A2_00973</td>
<td>polar_amino_acid_ABC_transporter</td>
<td>-1.54</td>
<td>1.91E-30</td>
</tr>
<tr>
<td>Elenta-A2_00729</td>
<td>extracellular_solute-binding_protein</td>
<td>-1.6</td>
<td>5.32E-16</td>
</tr>
<tr>
<td>Elenta-A2_00458</td>
<td>heat_shock_protein_Hsp20</td>
<td>-1.7</td>
<td>1.35E-13</td>
</tr>
<tr>
<td>Elenta-A2_01911</td>
<td>4Fe-4S_ferredoxin_iron-sulfur_binding_domain_protein</td>
<td>-1.71</td>
<td>2.48E-17</td>
</tr>
<tr>
<td>Elenta-A2_01480</td>
<td>22C5-didehydrogluconate_reductase</td>
<td>-1.79</td>
<td>8.27E-34</td>
</tr>
</tbody>
</table>
### Table 4.7 (Continued).

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_01940</td>
<td>Mg2_transporter_protein</td>
<td>-1.85</td>
<td>8.42E-35</td>
</tr>
<tr>
<td>Elenta-A2_02609</td>
<td>lipase_class_3</td>
<td>-1.9</td>
<td>6.13E-27</td>
</tr>
<tr>
<td>Elenta-A2_01586</td>
<td>membrane_protein_of_unknown_function</td>
<td>-1.91</td>
<td>4.52E-11</td>
</tr>
<tr>
<td>Elenta-A2_00505</td>
<td>hypothetical_protein</td>
<td>-1.99</td>
<td>2.65E-26</td>
</tr>
<tr>
<td>Elenta-A2_01113</td>
<td>hypothetical_protein</td>
<td>-2.02</td>
<td>1.42E-09</td>
</tr>
<tr>
<td>Elenta-A2_01910</td>
<td>molydopterin_dinucleotide-binding_region</td>
<td>-2.09</td>
<td>1.29E-56</td>
</tr>
<tr>
<td>Elenta-A2_01464</td>
<td>molydopterin_dinucleotide-binding_region</td>
<td>-2.11</td>
<td>6.01E-18</td>
</tr>
<tr>
<td>Elenta-A2_00508</td>
<td>Polysulphide_reductase_NrfD</td>
<td>-2.15</td>
<td>9.73E-25</td>
</tr>
<tr>
<td>Elenta-A2_01462</td>
<td>hypothetical_protein</td>
<td>-2.15</td>
<td>2.94E-10</td>
</tr>
<tr>
<td>Elenta-A2_00506</td>
<td>molybdopterin_oxidoreductase</td>
<td>-2.28</td>
<td>6.51E-37</td>
</tr>
<tr>
<td>Elenta-A2_00676</td>
<td>Radical_SAM_domain_protein</td>
<td>-2.56</td>
<td>3.98E-15</td>
</tr>
<tr>
<td>Elenta-A2_02199</td>
<td>acterioferritin</td>
<td>-2.6</td>
<td>4.10E-11</td>
</tr>
<tr>
<td>Elenta-A2_00507</td>
<td>4Fe-4S_ferredoxin_iron-sulfur_binding_domain_protein</td>
<td>-2.89</td>
<td>1.77E-11</td>
</tr>
</tbody>
</table>

### Table 4.8 Percent amino acid identity between the putative catechol dehydroxylases from *G. pamelaeae* 3C and *E. lenta* A2.

Ach stands for acetylene hydratase from *P. acetylenicus*. This protein is the biochemically characterized enzyme with the closest sequence similarity to the newly identified catechol dehydroxylases. Colors represent % identity, with identity going from low (blue) to high (red).

<table>
<thead>
<tr>
<th></th>
<th>Ach</th>
<th>Cldh</th>
<th>Dodh</th>
<th>Dadh</th>
<th>Cadh</th>
<th>Hcdh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ach</td>
<td>26</td>
<td>23.9</td>
<td>19.3</td>
<td>20.7</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Cldh</td>
<td>25.8</td>
<td>45.2</td>
<td>22</td>
<td>23.2</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Dodh</td>
<td>23.9</td>
<td>45</td>
<td>20.2</td>
<td>22</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Dadh</td>
<td>19.3</td>
<td>22</td>
<td>20.2</td>
<td></td>
<td>50.9</td>
<td>35.3</td>
</tr>
<tr>
<td>Cadh</td>
<td>20.7</td>
<td>23</td>
<td>22</td>
<td>50.9</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>Hcdh</td>
<td>15.7</td>
<td>17</td>
<td>15.7</td>
<td>35.3</td>
<td>39</td>
<td></td>
</tr>
</tbody>
</table>
To evaluate the involvement of a molybdenum enzyme in each dehydroxylation reaction, we cultured *E. lenta* A2 in the presence of tungstate (52). As with dopamine dehydroxylation, tungstate inhibited dehydroxylation of (+)-catechin and hydrocaffeic acid by *E. lenta* A2 without inhibiting growth in rich medium (BHI), suggesting these activities are indeed molybdenum dependent (Figures 4.25 and 4.26). Tungstate did not inhibit reduction of the benzyl ether in (+)-catechin, indicating this step is likely performed by a distinct enzyme (Figure 4.26). Finally, we found that the distribution of the genes encoding these the putative hydrocaffeic acid and (+)-catechin dehydroxylating enzymes across closely related *Eggerthella* strains correlated with metabolism of each substrate (Figure 4.27). For example, all *Eggerthella* strains except AN5LG harbored the putative hydrocaffeic acid dehydroxylase and could dehydroxylate this substrate. Similarly, carriage of the putative catechin dehydroxylase correlated with (+)-catechin metabolism, except for in the case of strain AB12#2, which did not encode for the enzyme but still had low metabolism (<10%) (Figure 4.27). This suggests that another enzyme might metabolize (+)-catechin in this strain. Overall, these data suggest that *Eggerthella* uses distinct molybdenum-dependent enzymes to dehydroxylate hydrocaffeic acid and (+)-catechin.
Figure 4.25 Tungstate does not inhibit *E. lenta* A2 growth in BHI medium.
*E. lenta* was grown anaerobically in BHI medium at 37 °C with 0.5 mM tungstate for 48 hours before growth was assessed. Bars represent the mean ± the SEM of three biological replicates. At 0.5 mM, the concentration at which tungstate inhibits catechol dehydroxylation by *E. lenta*, there is no effect on growth.
Figure 4.26 Impact of tungstate on catechol metabolism by E. lenta A2 and G. pamelaeae 3C.
A-C) Culturing in the presence of tungstate reveals the molybdenum dependence of catechol dehydroxylation by E. lenta A2. The strain was grown anaerobically with either 1 mM dopamine (A), hydrocaffeic acid (B), or (+)-catechin (C) and with either 0.5 mM tungstate (WO$_4^{2-}$) or molybdate (MoO$_4^{2-}$) for 48 hours 37 °C. Metabolites in culture supernatants were analyzed by LC-MS. Bars represent the mean ± the SEM of three biological replicates (bacterial cultures). The mass corresponding to m/z 275.0925 corresponds to the benzyl ether reduced, dehydroxylated catechin derivative. The experiment was performed twice. D) Tungstate does not inhibit benzyl ether reduction of (+)-catechin by E. lenta A2. E. lenta was grown anaerobically in BHI medium at 37 °C with 1 mM (+)-catechin and either molybdate (MoO$_4^{2-}$) or tungstate (WO$_4^{2-}$) (0.5 mM each) for 48 hours. Culture supernatants were analyzed by LC-MS/MS. Bars represent the mean ± the SEM mass peak area of the benzyl ether reduced catechin derivative (three biological replicates). Treatment of cultures with tungstate blocks dehydroxylation, thus leading to build-up of the benzyl ether reduced catechin derivative (m/z 291.0874). The experiment was performed twice. E) Tungstate does not inhibit DOPAC dehydroxylation by G. pamelaeae 3C. G. pamelaeae 3C was grown anaerobically with 1 mM DOPAC and either tungstate (WO$_4^{2-}$) or molybdate (MoO$_4^{2-}$) (0.5 mM each) for 48 hours 37 °C. The concentration of m-HPAA, the product of direct DOPAC dehydroxylation, in culture supernatants was analyzed by LC-MS. Bars represent the mean ± the SEM of three biological replicates (bacterial cultures).
Figure 4.27 Candidate catechol dehydroxylases and their associated metabolic activities are variably distributed across the gut Actinobacterial library used in our study. The tree represents the phylogeny of gut Actinobacterial strains adapted from (2). El = *Eggerthella lenta*, Es = *Eggerthella sinesis*, Ph = *Paraggerthella*, Gs = *Gordonibacter* sp., Gp = *Gordonibacter pamelaeae*. Squares represent gene presence/absence of select dehydroxylases across gut Actinobacterial strains (90% coverage, 75% amino acid identity cutoff, e-value=0). To evaluate catechol metabolism, individual strains were grown in triplicate in the presence of a single catechol substrate for 48 hours at 37 °C in BHI medium. Metabolism was assessed using LC-MS/MS. A red dot indicates that the mass peak of the dehydroxylated product was detected in cultures from this strain, while a black dot indicates lack of metabolism.
To biochemically validate one of our candidate dehydroxylases, we adapted the native purification protocol used for Dadh to fractionate the hydrocaffeic acid dehydroxylase activity from *E. lenta* A2 cell lysates. This anaerobic native protein preparation, which included ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, and size exclusion chromatography, yielded an active fraction that contained five major bands as assessed by SDS-PAGE (Figure 4.28) and quantitatively dehydroxylated hydrocaffeic acid into *m*-hydroxyphenylacetic acid under anaerobic conditions (Figure 4.29). We confirmed that the band with the apparent correct size (Figure 4.28) contained the proposed hydrocaffeic acid dehydroxylase (Hcdh, Elenta-A2_02815) using proteomics (Figure 4.30 and Table 4.9). The protein band also contained three other proteins (Table 4.9). One of these contaminants was a pyruvate dehydrogenase that was also observed in proteomics analyses of our Dadh purification (see Chapter 3 for details). These data link the newly identified *hcdh* gene to hydrocaffeic acid dehydroxylation, and, together with our discovery and verification of *dadh*, highlight the promise of using RNA-sequencing for enzyme discovery in *E. lenta*.

We performed an additional set of enzyme assays to evaluate the substrate scope of Hcdh. Consistent with our experiments in cell lysates (Figure 4.24), we observed dehydroxylation of hydrocaffeic acid and no reactivity towards dopamine, (+)-catechin, or DOPAC (Figure 4.31). Further evaluation of a subset of phenyl acid substrates revealed that Hcdh also metabolized caffeic acid, which differs from hydrocaffeic acid only in the degree of unsaturation. The free catechol was required for activity, as evidenced by the lack of activity towards ferulic acid, a molecule that differs from caffeic acid by the presence of a methoxy group at the *meta* position (Figure 4.31). Together, these data biochemically link the newly identified *hcdh* gene to hydrocaffeic acid dehydroxylation. The requirement of a catechol and the overall specificity of Hcdh is consistent
with the narrow substrate scope of Dadh, further supporting the proposal that different enzymes dehydroxylate distinct catechol substrates.

Figure 4.28 SDS-PAGE of partially purified hydrocaffeic acid dehydroxylase from *E. lenta* A2.

Ladder is the Precision Plus Protein™ All Blue Standards (first lane from the left), while the subsequent lane represents the combined hydrocaffeic acid-dehydroxylating fractions from the size exclusion column, the last chromatography step of the activity-based purification from *E. lenta* A2. The band highlighted with the red asterisk was cut out, submitted for proteomics, and was confirmed to contain the hydrocaffeic acid dehydroxylase (134 kDa predicted size with TAT signal sequence; 130.5 kDa predicted size without signal sequence).
Figure 4.29 In vitro activity of Hcdh-containing fractions purified from E. lenta A2.

EICs for detection of hydrocaffeic acid (MRM m/z 181->137) and m-hydroxyphenylpropionic acid (m-HPPA) (MRM m/z 165->121) after 26 hours of anaerobic incubation of enzyme preparation with 500 µM hydrocaffeic acid, 500 µM methyl viologen, and 1 mM sodium dithionite at room temperature. Peak heights show the relative intensity of each mass, and all chromatograms are shown on the same scale. The experiment was performed under anaerobic conditions unless otherwise indicated.
Figure 4.30 Peptide coverage of the hydrocaffeic acid dehydroxylase from *E. lenta* A2 from proteomics.

The band highlighted with a red asterisk in Figure 4.28 was cut out and subjected to proteomics. The sequence is the full hydrocaffeic acid dehydroxylase sequence (Uniprot accession #A0A369MIX7), and green indicates peptides identified in the proteomics experiment that were mapped to this sequence. As expected, there was no coverage at the beginning of the peptide sequence corresponding to the predicted TAT signal sequence, suggesting that it is cleaved during protein maturation. Whereas the predicted size of the non-processed hydrocaffeic acid dehydroxylase is 133.9 kDa, the mature peptide that has undergone TAT signal sequence cleavage is expected to have a size of 130.5 kDa.

Table 4.9 Proteomics identification of proteins in the band of interest during the activity-based native purification of the hydrocaffeic acid dehydroxylase from *E. lenta* A2.

The band highlighted with a red asterisk in Figure 4.28 was cut out and subjected to proteomics. This band was confirmed to contain the catalytic subunit of the hydrocaffeic acid dehydroxylase (highlighted in red). The predicted size of 133.9 kDa (a) is for the full peptide containing the Twin Arginine Translocation (TAT) signal sequence. The predicted size of 130.5 kDa (b) is for the processed, mature peptide where the TAT signal sequence has been removed, as is suggested from the coverage map in Figure 4.30.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>Coverage</th>
<th># AAs</th>
<th>MW [kDa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>237-Elenta-A2_02071</td>
<td>Pyruvate ferredoxin/flavodoxin oxidoreductase</td>
<td>82</td>
<td>1180</td>
<td>127.6</td>
</tr>
<tr>
<td>237-Elenta-A2_02815</td>
<td>Molybdopterin dinucleotide binding region</td>
<td>61</td>
<td>1193</td>
<td>133.9(a)/130.5(b)</td>
</tr>
<tr>
<td>237-Elenta-A2_01884</td>
<td>Single-stranded-DNA-specific exonuclease</td>
<td>54</td>
<td>1122</td>
<td>120.5</td>
</tr>
<tr>
<td>237-Elenta-A2_01074</td>
<td>UvrD/REP helicase</td>
<td>41</td>
<td>1165</td>
<td>123</td>
</tr>
</tbody>
</table>
Figure 4.31 *E. lenta* A2 Hcdh displays a narrow substrate scope and requires the hydrocaffeic acid core scaffold for metabolism.

A) Structures of hydrocaffeic acid and its phenylcarboxylic acid analogs evaluated with Hcdh. B) Hcdh activity assays. The enzyme preparation used in Figure 4.29 was diluted 1:5 in 50 mM Tris pH 8 buffer and was incubated with 500 µM catechol substrate, 500 µM methyl viologen, and 1 mM sodium dithionite at room temperature for 26 hours under anaerobic conditions. The enzyme reactions were analyzed by LC-MS/MS. Bars represent the mean ± the SEM of three independent enzyme reactions.
4.2.7 Discovery of a candidate DOPAC dehydroxylase.

We next sought to identify the enzyme responsible for DOPAC dehydroxylation in *G. pamelaeeae* 3C. We first profiled the timing of DOPAC dehydroxylation in cultures of *G. pamelaeeae* 3C in which the compound was either present from beginning or was spiked in during the exponential phase (Figure 4.32). Based on this experiment, we performed two separate RNA-sequencing experiments. First, we added DOPAC to *G. pamelaeeae* 3C cultures at mid-exponential phase and harvested cells after 3 hours of induction when the cultures had reached early stationary phase. In this experiment, *G. pamelaeeae* 3C upregulated 99 different genes, including four distinct molybdenum-dependent enzyme-encoding genes (Table 4.10). One of these genes (C1877_13905) was among the highest upregulated genes across the dataset (>1700-fold induced). To further explore the association between this gene and DOPAC dehydroxylation, we repeated the RNA-seq experiment, growing *G. pamelaeeae* 3C in the presence of DOPAC from the time of inoculation and then harvesting cells in exponential phase as soon we could detect metabolism (12 hours of growth). We speculated that changing the conditions in this manner would allow us to more clearly observe the genes that were expressed early in response to DOPAC. We also anticipated that only genes expressed across both experimental setups would be reasonable DOPAC dehydroxylases. In this experiment, the same molybdenum-dependent enzyme-encoding gene (C1877_13905) that was highly upregulated in our first experiment (Table 4.11) was among the highest upregulated genes. The only two other molybdenum-dependent enzymes induced in this experiment (which were also upregulated in the first experimental setup) were induced at an order of magnitude lower levels (2-fold induced). Based on these results, we propose that the molybdenum-dependent enzyme encoded by C1877_13905 is a candidate DOPAC dehydroxylase.
This assignment is also supported by comparative genomics. First, carriage of C1877_13905 (named dodh) correlated with DOPAC dehydroxylation among members of our gut Actinobacterial library (Figure 4.27). This activity was only present in the three non-Eggerthella strains in the library, including the two Gordonibacter and one Paraeggerthella species. Consistent with our lysate assays, those organisms harboring this gene also had activity towards hydrocaffeic acid, which could potentially explain the pattern of hydrocaffeic acid metabolism across the gut Actinobacterial library (Figure 4.27). Finally, the biochemically characterized enzyme most similar to the candidate DOPAC dehydroxylase is Dadh (20% amino acid ID) (Table 4.8). Interestingly, unlike with the Eggerthella dehydroxylases, tungstate did not inhibit dehydroxylation of DOPAC by G. pamelaeae (Figure 4.26). This may be explained if the dehydroxylating enzyme can use both molybdenum and tungsten for catalysis, as is seen in certain closely related enzymes (53).

During the time we conducted our study, our collaborators in the Turnbaugh lab at UCSF identified a candidate Gordonibacter dehydroxylase that metabolizes the lignan dmSECO from flaxseed (19). To locate this enzyme, they also compared differentially expressed transcripts between G. pamelaeae 3C grown with dmSECO and a vehicle control. They identified a set of highly upregulated genes, including a single gene encoding for a molybdenum-dependent enzyme that was upregulated >2000-fold. Named cldh, this gene was one of 13 differentially expressed genes that were unique to metabolizing Gordonibacter strains but were not present in other closely related non-metabolizing Actinobacteria. Interestingly, a subset of genes (not including the molybdenum-dependent enzymes) expressed in response to DOPAC were also upregulated in response to dmSECO, suggesting catechols may elicit shared transcriptional responses in Gordonibacter (19). Elizabeth Bess (now a professor at UC Irvine), the first author of this paper,
also mentioned that tungstate did not inhibit dmSECO dehydroxylation in this organism, similar to our observations with DOPAC metabolism (Figure 4.26). While the authors did not biochemically confirm the activity of Cldh, the high percent amino acid identity (45%) of this enzyme to Dodh and its relationship to Dadh, Hcdh, and Cadh strongly supports its involvement in the dehydroxylation of dietary lignans (Table 4.8).

![Figure 4.32 Time course of DOPAC dehydroxylation by *G. pamelaeeae* 3C.](image-url)

This strain was grown in BHI medium containing 10 mM formate with and without 0.5 mM DOPAC at 37 °C. Growth and metabolism were measured periodically. Metabolism was assessed using the catechol colorimetric method. Bars represent the mean ± the SEM from three biological replicates, e.g. bacterial cultures. A) Growth and B) metabolism when DOPAC was added at the time of inoculation, as indicated by the black arrow. C) Growth and D) metabolism when DOPAC was added to exponentially growing cells, as indicated by the black arrow.
Table 4.10 Genes differentially expressed upon exposure of *G. pamelaee* 3C to DOPAC when the catechol is added during exponential phase

The data represent the genes that are differentially expressed in response to exposure of *G. pamelaee* 3C to 0.5 mM DOPAC relative to vehicle (>2-fold difference, FDR<0.1). The catalytic subunit of the putative DOPAC dehydroxylase (*dodh*), as well as its predicted 4Fe-4S partner, are highlighted in red. The data are from n=3 cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1877_00905</td>
<td>Radical_SAM_protein</td>
<td>11.21</td>
<td>5.39E-80</td>
</tr>
<tr>
<td>C1877_13890</td>
<td>MFS_transporter</td>
<td>11.16</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00920</td>
<td>nitrate_ABC_transporter_substrate-binding_protein</td>
<td>10.96</td>
<td>1.32E-49</td>
</tr>
<tr>
<td>C1877_13905</td>
<td>Dehydrogenase (<em>dodh</em>)</td>
<td>10.81</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13900</td>
<td>Oxidoreductase (<em>dodh</em> 4Fe-4S partner)</td>
<td>10.63</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00910</td>
<td>ABC_transporter_permease</td>
<td>10.28</td>
<td>1.31E-104</td>
</tr>
<tr>
<td>C1877_13910</td>
<td>Hypothetical_protein</td>
<td>10.27</td>
<td>1.64E-171</td>
</tr>
<tr>
<td>C1877_13895</td>
<td>Hypothetical_protein</td>
<td>10.1</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00915</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>9.66</td>
<td>1.72E-116</td>
</tr>
<tr>
<td>C1877_00925</td>
<td>BC_transporter_permease</td>
<td>9.25</td>
<td>7.05E-46</td>
</tr>
<tr>
<td>C1877_07205</td>
<td>GTPase_G3E_family</td>
<td>8.99</td>
<td>1.57E-271</td>
</tr>
<tr>
<td>C1877_13865</td>
<td>rimF:ribosomal-protein-alanine_N-acetyltransferase</td>
<td>8.92</td>
<td>2.22E-252</td>
</tr>
<tr>
<td>C1877_13885</td>
<td>hypothetical_protein</td>
<td>8.9</td>
<td>6.02E-288</td>
</tr>
<tr>
<td>C1877_07570</td>
<td>hypothetical_protein</td>
<td>8.78</td>
<td>1.07E-254</td>
</tr>
<tr>
<td>C1877_04445</td>
<td>sodium:solute_symporter</td>
<td>8.58</td>
<td>0</td>
</tr>
<tr>
<td>C1877_07565</td>
<td>dehydrogenase</td>
<td>8.32</td>
<td>5.42E-93</td>
</tr>
<tr>
<td>C1877_13930</td>
<td>hypothetical_protein</td>
<td>8.28</td>
<td>4.42E-32</td>
</tr>
<tr>
<td>C1877_13870</td>
<td>hypothetical_protein</td>
<td>8.04</td>
<td>1.77E-70</td>
</tr>
<tr>
<td>C1877_07200</td>
<td>DUF3641_domain-containing_protein</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13875</td>
<td>hypothetical_protein</td>
<td>7.85</td>
<td>0</td>
</tr>
<tr>
<td>C1877_07185</td>
<td>ABC_transporter_substrate-binding_protein</td>
<td>7.85</td>
<td>6.85E-131</td>
</tr>
<tr>
<td>C1877_07195</td>
<td>ABC_transporter_permease</td>
<td>7.74</td>
<td>1.21E-235</td>
</tr>
<tr>
<td>C1877_07180</td>
<td>DUF1847_domain-containing_protein</td>
<td>7.59</td>
<td>3.20E-64</td>
</tr>
<tr>
<td>C1877_07580</td>
<td>MFS_transporter</td>
<td>7.55</td>
<td>0</td>
</tr>
<tr>
<td>C1877_07210</td>
<td>methyltransferase</td>
<td>7.34</td>
<td>4.27E-300</td>
</tr>
<tr>
<td>C1877_07190</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>7.23</td>
<td>7.82E-100</td>
</tr>
<tr>
<td>C1877_13880</td>
<td>hypothetical_protein</td>
<td>7.04</td>
<td>1.69E-91</td>
</tr>
<tr>
<td>C1877_09915</td>
<td>radical_SAM_protein</td>
<td>6.85</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00930</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>6.44</td>
<td>1.27E-103</td>
</tr>
</tbody>
</table>
Table 4.10 (Continued).

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1877_09910</td>
<td>hypothetical_protein</td>
<td>6.36</td>
<td>0</td>
</tr>
<tr>
<td>C1877_04725</td>
<td>hypothetical_protein</td>
<td>5.95</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00940</td>
<td>hypothetical_protein</td>
<td>5.89</td>
<td>1.85E-133</td>
</tr>
<tr>
<td>C1877_13835</td>
<td>DUF3641_domain-containing_protein</td>
<td>5.56</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13840</td>
<td>4-carboxymuconolactone_decarboxylase</td>
<td>5.32</td>
<td>0</td>
</tr>
<tr>
<td>C1877_08360</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>5.24</td>
<td>1.80E-24</td>
</tr>
<tr>
<td>C1877_07315</td>
<td>hypothetical_protein</td>
<td>5.22</td>
<td>1.59E-220</td>
</tr>
<tr>
<td>C1877_13860</td>
<td>hypothetical_protein</td>
<td>5.14</td>
<td>0</td>
</tr>
<tr>
<td>C1877_07560</td>
<td>oxidoreductase</td>
<td>5.03</td>
<td>4.60E-210</td>
</tr>
<tr>
<td>C1877_13785</td>
<td>4Fe-4S_ferredoxin</td>
<td>4.89</td>
<td>2.47E-20</td>
</tr>
<tr>
<td>C1877_02200</td>
<td>ATP-binding_protein</td>
<td>4.81</td>
<td>0.020782133</td>
</tr>
<tr>
<td>C1877_08370</td>
<td>16S_rRNA_Adenine_N6-dimethyltransferase_RsmA</td>
<td>4.8</td>
<td>7.32E-09</td>
</tr>
<tr>
<td>C1877_08365</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>4.78</td>
<td>3.06E-14</td>
</tr>
<tr>
<td>C1877_13785</td>
<td>4Fe-4S_ferredoxin</td>
<td>4.78</td>
<td>4.68E-300</td>
</tr>
<tr>
<td>C1877_13790</td>
<td>dehydrogenase</td>
<td>4.61</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13780</td>
<td>hypothetical_protein</td>
<td>4.53</td>
<td>2.22E-238</td>
</tr>
<tr>
<td>C1877_08355</td>
<td>energy-coupling_factor_transporter</td>
<td>4.25</td>
<td>4.16E-21</td>
</tr>
<tr>
<td>C1877_04450</td>
<td>hypothetical_protein</td>
<td>4.02</td>
<td>1.11E-292</td>
</tr>
<tr>
<td>C1877_01125</td>
<td>phenylacetate--CoA_ligase</td>
<td>3.97</td>
<td>2.44E-97</td>
</tr>
<tr>
<td>C1877_05190</td>
<td>hypothetical_protein</td>
<td>3.93</td>
<td>0.036362059</td>
</tr>
<tr>
<td>C1877_03100</td>
<td>hypothetical_protein</td>
<td>3.82</td>
<td>5.89E-68</td>
</tr>
<tr>
<td>C1877_14220</td>
<td>hypothetical_protein</td>
<td>3.74</td>
<td>0.043112489</td>
</tr>
<tr>
<td>C1877_12815</td>
<td>arylamine_N-acetyltransferase</td>
<td>3.69</td>
<td>0</td>
</tr>
<tr>
<td>C1877_14165</td>
<td>conjugal_transfer_protein_TraC</td>
<td>3.65</td>
<td>0.045993369</td>
</tr>
<tr>
<td>C1877_00945</td>
<td>hypothetical_protein</td>
<td>3.6</td>
<td>0.001453692</td>
</tr>
<tr>
<td>C1877_13920</td>
<td>LuxR_family_transcriptional_regulator</td>
<td>3.42</td>
<td>9.77E-108</td>
</tr>
<tr>
<td>C1877_13850</td>
<td>hypothetical_protein</td>
<td>3.41</td>
<td>1.34E-201</td>
</tr>
<tr>
<td>C1877_13915</td>
<td>hypothetical_protein</td>
<td>3.39</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13925</td>
<td>hypothetical_protein</td>
<td>3.18</td>
<td>2.86E-187</td>
</tr>
<tr>
<td>C1877_01120</td>
<td>hypothetical_protein</td>
<td>3.01</td>
<td>1.50E-263</td>
</tr>
<tr>
<td>C1877_13775</td>
<td>hypothetical_protein</td>
<td>3</td>
<td>1.80E-127</td>
</tr>
<tr>
<td>C1877_13770</td>
<td>dehydrogenase</td>
<td>2.94</td>
<td>0</td>
</tr>
<tr>
<td>C1877_07550</td>
<td>hypothetical_protein</td>
<td>2.92</td>
<td>3.78E-119</td>
</tr>
</tbody>
</table>
Table 4.10 (Continued).

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1877_13970</td>
<td>amidohydrolase</td>
<td>2.74</td>
<td>6.24E-178</td>
</tr>
<tr>
<td>C1877_13765</td>
<td>DUF2064_domain-containing_protein</td>
<td>2.53</td>
<td>9.85E-74</td>
</tr>
<tr>
<td>C1877_03335</td>
<td>TrpB-like_pyridoxal_phosphate-dependent_enzyme</td>
<td>2.52</td>
<td>7.14E-11</td>
</tr>
<tr>
<td>C1877_05595</td>
<td>hypothetical_protein</td>
<td>2.42</td>
<td>0.070104052</td>
</tr>
<tr>
<td>C1877_09905</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>2.36</td>
<td>4.47E-128</td>
</tr>
<tr>
<td>C1877_07555</td>
<td>hypothetical_protein</td>
<td>2.17</td>
<td>1.47E-56</td>
</tr>
<tr>
<td>C1877_06120</td>
<td>hypothetical_protein</td>
<td>2.05</td>
<td>5.92E-64</td>
</tr>
<tr>
<td>C1877_07555</td>
<td>4Fe-4S-binding_protein</td>
<td>2</td>
<td>2.31E-93</td>
</tr>
<tr>
<td>C1877_07585</td>
<td>helix-turn-helix_transcriptional_regulator</td>
<td>1.96</td>
<td>3.32E-60</td>
</tr>
<tr>
<td>C1877_09900</td>
<td>hypothetical_protein</td>
<td>1.95</td>
<td>7.02E-82</td>
</tr>
<tr>
<td>C1877_13760</td>
<td>radical_SAM_protein</td>
<td>1.93</td>
<td>1.44E-163</td>
</tr>
<tr>
<td>C1877_13750</td>
<td>hypothetical_protein</td>
<td>1.88</td>
<td>2.68E-59</td>
</tr>
<tr>
<td>C1877_13230</td>
<td>formate_dehydrogenase</td>
<td>1.87</td>
<td>0.003539835</td>
</tr>
<tr>
<td>C1877_00935</td>
<td>hypothetical_protein</td>
<td>1.79</td>
<td>4.10E-22</td>
</tr>
<tr>
<td>C1877_04455</td>
<td>endoglucanase</td>
<td>1.74</td>
<td>3.14E-05</td>
</tr>
<tr>
<td>C1877_13855</td>
<td>hypothetical_protein</td>
<td>1.73</td>
<td>1.73E-73</td>
</tr>
<tr>
<td>C1877_00900</td>
<td>hypothetical_protein</td>
<td>1.72</td>
<td>0.010146811</td>
</tr>
<tr>
<td>C1877_07535</td>
<td>hypothetical_protein</td>
<td>1.71</td>
<td>7.39E-45</td>
</tr>
<tr>
<td>C1877_07295</td>
<td>MFS_transporter</td>
<td>1.66</td>
<td>0.01714762</td>
</tr>
<tr>
<td>C1877_14955</td>
<td>hypothetical_protein</td>
<td>1.63</td>
<td>8.11E-66</td>
</tr>
<tr>
<td>C1877_07310:</td>
<td>hypothetical_protein</td>
<td>1.59</td>
<td>4.20E-07</td>
</tr>
<tr>
<td>C1877_02570</td>
<td>MFS_transporter</td>
<td>1.53</td>
<td>1.11E-34</td>
</tr>
<tr>
<td>C1877_14780:</td>
<td>site-2_protease_family_protein</td>
<td>1.5</td>
<td>0.000440981</td>
</tr>
<tr>
<td>C1877_02990</td>
<td>DUF488_domain-containing_protein:</td>
<td>1.33</td>
<td>0.04199253</td>
</tr>
<tr>
<td>C1877_07300</td>
<td>helix-turn-helix_transcriptional_regulator</td>
<td>1.28</td>
<td>7.21E-08</td>
</tr>
<tr>
<td>C1877_04550:</td>
<td>hypothetical_protein</td>
<td>1.25</td>
<td>1.64E-16</td>
</tr>
<tr>
<td>C1877_02800</td>
<td>MFS_transporter</td>
<td>1.22</td>
<td>0.045193782</td>
</tr>
<tr>
<td>C1877_12325</td>
<td>AraC_family_transcriptional_regulator</td>
<td>1.13</td>
<td>0.027997941</td>
</tr>
<tr>
<td>C1877_03030</td>
<td>DUF4405_domain-containing_protein</td>
<td>1.09</td>
<td>5.62E-16</td>
</tr>
<tr>
<td>C1877_00895</td>
<td>MFS_transporter</td>
<td>1.09</td>
<td>4.56E-05</td>
</tr>
<tr>
<td>C1877_07575</td>
<td>hypothetical_protein</td>
<td>1.08</td>
<td>2.59E-27</td>
</tr>
<tr>
<td>C1877_09415</td>
<td>GNAT_family_N-acetyltransferase</td>
<td>1.08</td>
<td>9.38E-18</td>
</tr>
<tr>
<td>Locus tag</td>
<td>Annotation</td>
<td>log2FoldChange</td>
<td>p-value (FDR&lt;0.1)</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>C1877_03295:</td>
<td>hypothetical_protein</td>
<td>1.08</td>
<td>0.000322218</td>
</tr>
<tr>
<td>C1877_03035</td>
<td>nickel-dependent_hydrogenase_large_subunit</td>
<td>1.03</td>
<td>8.67E-19</td>
</tr>
<tr>
<td>C1877_09205</td>
<td>amino_acid_ABC_transporter_permease</td>
<td>1.01</td>
<td>0.026568145</td>
</tr>
<tr>
<td>C1877_01130</td>
<td>LysR_family_transcriptional_regulator</td>
<td>1.01</td>
<td>1.73E-14</td>
</tr>
<tr>
<td>C1877_14775</td>
<td>3-oxoacyl-ACP_reductase</td>
<td>1</td>
<td>0.004320782</td>
</tr>
<tr>
<td>C1877_06060</td>
<td>hypothetical_protein</td>
<td>-1.08</td>
<td>7.29E-12</td>
</tr>
<tr>
<td>C1877_06730</td>
<td>polysulfide_reductase</td>
<td>-1.08</td>
<td>3.98E-06</td>
</tr>
<tr>
<td>C1877_06720</td>
<td>hypothetical_protein</td>
<td>-1.09</td>
<td>7.77E-27</td>
</tr>
<tr>
<td>C1877_10640</td>
<td>hypothetical_protein</td>
<td>-1.13</td>
<td>4.05E-17</td>
</tr>
<tr>
<td>C1877_10720</td>
<td>LysR_family_transcriptional_regulator</td>
<td>-1.16</td>
<td>4.49E-22</td>
</tr>
<tr>
<td>C1877_06070:</td>
<td>hypothetical_protein</td>
<td>-1.17</td>
<td>1.93E-10</td>
</tr>
<tr>
<td>C1877_10665</td>
<td>putative_sulfate_exporter</td>
<td>-1.29</td>
<td>1.91E-05</td>
</tr>
<tr>
<td>C1877_10635</td>
<td>hypothetical_protein</td>
<td>-1.29</td>
<td>2.87E-19</td>
</tr>
<tr>
<td>C1877_10715</td>
<td>nitrate_reductase_gamma_subunit</td>
<td>-1.35</td>
<td>1.71E-16</td>
</tr>
<tr>
<td>C1877_10690</td>
<td>porphobilinogen_synthase</td>
<td>-1.38</td>
<td>1.14E-10</td>
</tr>
<tr>
<td>C1877_10685</td>
<td>hemL:glutamate-l-semialdehyde-aminomutase</td>
<td>-1.42</td>
<td>7.29E-12</td>
</tr>
<tr>
<td>C1877_10660</td>
<td>hypothetical_protein</td>
<td>-1.45</td>
<td>2.18E-05</td>
</tr>
<tr>
<td>C1877_10605</td>
<td>sulfurtransferase_TusE</td>
<td>-1.51</td>
<td>2.66E-14</td>
</tr>
<tr>
<td>C1877_10630</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.53</td>
<td>4.72E-23</td>
</tr>
<tr>
<td>C1877_10600</td>
<td>cobyrinic_acid_-diamide_synthase</td>
<td>-1.56</td>
<td>5.62E-16</td>
</tr>
<tr>
<td>C1877_10705</td>
<td>glutamyl-tRNA_reductase</td>
<td>-1.6</td>
<td>8.69E-18</td>
</tr>
<tr>
<td>C1877_10700</td>
<td>hydroxymethylbilane_synthase</td>
<td>-1.61</td>
<td>3.95E-15</td>
</tr>
<tr>
<td>C1877_10695</td>
<td>cobA:uroporphyrinogen-III_C-methyltransferase</td>
<td>-1.63</td>
<td>4.61E-10</td>
</tr>
<tr>
<td>C1877_10610</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.63</td>
<td>1.98E-27</td>
</tr>
<tr>
<td>C1877_10710</td>
<td>Fe-S-binding_protein</td>
<td>-1.64</td>
<td>1.89E-17</td>
</tr>
<tr>
<td>C1877_05025</td>
<td>dehydrogenase</td>
<td>-1.68</td>
<td>1.09E-05</td>
</tr>
<tr>
<td>C1877_10625</td>
<td>Fe-S-binding_protein</td>
<td>-1.68</td>
<td>1.46E-26</td>
</tr>
<tr>
<td>C1877_10615</td>
<td>4Fe-4S_dicluster_domain-containing_protein</td>
<td>-1.71</td>
<td>1.62E-31</td>
</tr>
<tr>
<td>C1877_10620</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.77</td>
<td>3.21E-26</td>
</tr>
<tr>
<td>C1877_07910</td>
<td>hypothetical_protein</td>
<td>-2.12</td>
<td>0.000850235</td>
</tr>
<tr>
<td>C1877_06075</td>
<td>IS1_family_transposase</td>
<td>-2.16</td>
<td>0.051755192</td>
</tr>
<tr>
<td>C1877_05990</td>
<td>dehydrogenase</td>
<td>-2.2</td>
<td>0.068743035</td>
</tr>
<tr>
<td>C1877_06180</td>
<td>DUF1269_domain-containing_protein</td>
<td>-2.66</td>
<td>0.072663541</td>
</tr>
<tr>
<td>Locus tag</td>
<td>Annotation</td>
<td>log2FoldChange</td>
<td>p-value (FDR&lt;0.1)</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>C1877_02225</td>
<td>hypothetical_protein</td>
<td>-2.68</td>
<td>0.040089976</td>
</tr>
<tr>
<td>C1877_05995</td>
<td>oxidoreductase</td>
<td>-3.4</td>
<td>0.062311297</td>
</tr>
<tr>
<td>C1877_10895</td>
<td>MFS_transporter</td>
<td>-3.73</td>
<td>0.026568145</td>
</tr>
<tr>
<td>C1877_14300</td>
<td>hypothetical_protein</td>
<td>-4.12</td>
<td>0.076180484</td>
</tr>
<tr>
<td>C1877_02470</td>
<td>transporter</td>
<td>-5.02</td>
<td>0.002337819</td>
</tr>
</tbody>
</table>
Table 4.11 Genes differentially expressed upon exposure of *G. pamelaeae* 3C to DOPAC when the catechol is added at the time of inoculation.
The data represent the genes that are differentially expressed in response to exposure of *G. pamelaeae* 3C to 0.5 mM DOPAC relative to vehicle (>2-fold difference, FDR<0.1). The catalytic subunit of the putative DOPAC dehydroxylase (*dodh*), as well as its predicted 4Fe-4S partner, are highlighted in red. The data are from n=3 cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1877_00910</td>
<td>ABC_transporter_permease</td>
<td>8.05</td>
<td>4.65E-114</td>
</tr>
<tr>
<td>C1877_00920</td>
<td>nitrate_ABC_transporter_substrate-binding_protein</td>
<td>7.78</td>
<td>1.13E-160</td>
</tr>
<tr>
<td>C1877_13870</td>
<td>hypothetical_protein</td>
<td>7.42</td>
<td>3.46E-42</td>
</tr>
<tr>
<td>C1877_13910</td>
<td>hypothetical_protein</td>
<td>7.33</td>
<td>3.64E-88</td>
</tr>
<tr>
<td>C1877_07205</td>
<td>GTPase_G3E_family</td>
<td>7.23</td>
<td>1.53E-141</td>
</tr>
<tr>
<td>C1877_13890</td>
<td>MFS_transporter</td>
<td>7.22</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13905</td>
<td>Dehydrogenase (<em>dodh</em>)</td>
<td>7.2</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00915</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>7.18</td>
<td>3.24E-140</td>
</tr>
<tr>
<td>C1877_13900</td>
<td>Oxidoreductase (<em>dodh</em> 4Fe-4S partner)</td>
<td>7.11</td>
<td>0</td>
</tr>
<tr>
<td>C1877_13895</td>
<td>hypothetical_protein</td>
<td>7.07</td>
<td>0</td>
</tr>
<tr>
<td>C1877_00925</td>
<td>ABC_transporter_permease</td>
<td>7.05</td>
<td>1.11E-34</td>
</tr>
<tr>
<td>C1877_13930</td>
<td>hypothetical_protein</td>
<td>6.73</td>
<td>7.65E-19</td>
</tr>
<tr>
<td>C1877_13875</td>
<td>hypothetical_protein</td>
<td>6.56</td>
<td>1.51E-189</td>
</tr>
<tr>
<td>C1877_00905</td>
<td>radical_SAM_protein</td>
<td>6.53</td>
<td>1.31E-59</td>
</tr>
<tr>
<td>C1877_07190</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>6.44</td>
<td>3.33E-68</td>
</tr>
<tr>
<td>C1877_13865</td>
<td>ribosomal-protein-alanine_N-acetyltransferase</td>
<td>6.36</td>
<td>2.98E-44</td>
</tr>
<tr>
<td>C1877_07200</td>
<td>DUF3641_domain-containing_protein</td>
<td>6.12</td>
<td>7.13E-248</td>
</tr>
<tr>
<td>C1877_07195</td>
<td>ABC_transporter_permease</td>
<td>6.12</td>
<td>9.85E-146</td>
</tr>
<tr>
<td>C1877_07185</td>
<td>ABC_transporter_substrate-binding_protein</td>
<td>6.02</td>
<td>2.87E-155</td>
</tr>
<tr>
<td>C1877_07180</td>
<td>DUF1847_domain-containing_protein</td>
<td>6.01</td>
<td>2.12E-97</td>
</tr>
<tr>
<td>C1877_13885</td>
<td>hypothetical_protein</td>
<td>5.77</td>
<td>9.05E-135</td>
</tr>
<tr>
<td>C1877_07210</td>
<td>methyltransferase</td>
<td>5.43</td>
<td>4.47E-140</td>
</tr>
<tr>
<td>C1877_04445</td>
<td>sodium:solute_symporter</td>
<td>5.03</td>
<td>4.08E-263</td>
</tr>
<tr>
<td>C1877_13880</td>
<td>hypothetical_protein</td>
<td>4.16</td>
<td>1.04E-25</td>
</tr>
<tr>
<td>C1877_00930</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>3.88</td>
<td>4.79E-147</td>
</tr>
<tr>
<td>C1877_13835</td>
<td>DUF3641_domain-containing_protein</td>
<td>3.8</td>
<td>2.70E-254</td>
</tr>
<tr>
<td>C1877_0731</td>
<td>hypothetical_protein</td>
<td>3.7</td>
<td>3.07E-100</td>
</tr>
<tr>
<td>Locus tag</td>
<td>Annotation</td>
<td>log2FoldChange</td>
<td>p-value (FDR&lt;0.1)</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>C1877_00940</td>
<td>hypothetical_protein</td>
<td>3.7</td>
<td>2.50E-59</td>
</tr>
<tr>
<td>C1877_13840</td>
<td>4-carboxymuconolactone_decarboxylase</td>
<td>3.42</td>
<td>3.87E-196</td>
</tr>
<tr>
<td>C1877_09910</td>
<td>hypothetical_protein</td>
<td>3.36</td>
<td>3.42E-41</td>
</tr>
<tr>
<td>C1877_09915</td>
<td>hypothetical_protein</td>
<td>3.27</td>
<td>2.53E-244</td>
</tr>
<tr>
<td>C1877_09915:</td>
<td>radical_SAM_protein</td>
<td>2.68</td>
<td>6.77E-208</td>
</tr>
<tr>
<td>C1877_00935</td>
<td>hypothetical_protein</td>
<td>2.16</td>
<td>3.97E-38</td>
</tr>
<tr>
<td>C1877_04450</td>
<td>hypothetical_protein</td>
<td>1.81</td>
<td>1.44E-61</td>
</tr>
<tr>
<td>C1877_09900:</td>
<td>hypothetical_protein</td>
<td>1.67</td>
<td>2.49E-52</td>
</tr>
<tr>
<td>C1877_09905</td>
<td>ABC_transporter_ATP-binding_protein</td>
<td>1.6</td>
<td>5.09E-57</td>
</tr>
<tr>
<td>C1877_13860</td>
<td>hypothetical_protein</td>
<td>1.6</td>
<td>9.49E-24</td>
</tr>
<tr>
<td>C1877_14955</td>
<td>hypothetical_protein</td>
<td>1.41</td>
<td>2.20E-48</td>
</tr>
<tr>
<td>C1877_13775</td>
<td>hypothetical_protein</td>
<td>1.32</td>
<td>1.58E-25</td>
</tr>
<tr>
<td>C1877_13970</td>
<td>amidohydrolase</td>
<td>1.29</td>
<td>6.59E-41</td>
</tr>
<tr>
<td>C1877_13785</td>
<td>4Fe-4S_ferredoxin</td>
<td>1.27</td>
<td>4.40E-25</td>
</tr>
<tr>
<td>C1877_13790</td>
<td>dehydrogenase</td>
<td>1.23</td>
<td>1.36E-34</td>
</tr>
<tr>
<td>C1877_13780:</td>
<td>hypothetical_protein</td>
<td>1.21</td>
<td>4.91E-18</td>
</tr>
<tr>
<td>C1877_13750</td>
<td>hypothetical_protein</td>
<td>1.2</td>
<td>8.36E-32</td>
</tr>
<tr>
<td>C1877_13925</td>
<td>hypothetical_protein</td>
<td>1.2</td>
<td>3.24E-28</td>
</tr>
<tr>
<td>C1877_13770:</td>
<td>dehydrogenase</td>
<td>1.18</td>
<td>3.12E-61</td>
</tr>
<tr>
<td>C1877_13915:</td>
<td>hypothetical_protein</td>
<td>1.18</td>
<td>5.70E-44</td>
</tr>
<tr>
<td>C1877_13755</td>
<td>4Fe-4S-binding_protein</td>
<td>1.07</td>
<td>5.38E-33</td>
</tr>
<tr>
<td>C1877_10245</td>
<td>hypothetical_protein</td>
<td>-1.14</td>
<td>0.000531507</td>
</tr>
<tr>
<td>C1877_10635</td>
<td>hypothetical_protein</td>
<td>-1.33</td>
<td>1.41E-19</td>
</tr>
<tr>
<td>C1877_10625</td>
<td>Fe-S-binding_protein</td>
<td>-1.35</td>
<td>2.25E-16</td>
</tr>
<tr>
<td>C1877_10640</td>
<td>hypothetical_protein</td>
<td>-1.37</td>
<td>1.64E-23</td>
</tr>
<tr>
<td>C1877_10610</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.4</td>
<td>4.17E-19</td>
</tr>
<tr>
<td>C1877_10620</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.42</td>
<td>6.01E-15</td>
</tr>
<tr>
<td>C1877_10630</td>
<td>menaquinol_oxidoreductase</td>
<td>-1.45</td>
<td>7.21E-20</td>
</tr>
<tr>
<td>C1877_10710</td>
<td>Fe-S-binding_protein</td>
<td>-1.48</td>
<td>1.09E-13</td>
</tr>
<tr>
<td>C1877_10600</td>
<td>cobyrinic_acid_diamide_synthase</td>
<td>-1.51</td>
<td>2.26E-14</td>
</tr>
<tr>
<td>C1877_10615</td>
<td>4Fe-4S_dicluster_domain-containing_protein</td>
<td>-1.51</td>
<td>1.89E-22</td>
</tr>
<tr>
<td>C1877_10705</td>
<td>glutamyl-tRNA_reductase:</td>
<td>-1.55</td>
<td>4.09E-15</td>
</tr>
</tbody>
</table>
Table 4.1 (Continued).

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1877_10715</td>
<td>nitrate_reductase_gamma_subunit</td>
<td>-1.56</td>
<td>3.01E-20</td>
</tr>
<tr>
<td>C1877_02335</td>
<td>ferrous_iron_transporter_B</td>
<td>-1.57</td>
<td>1.60E-06</td>
</tr>
<tr>
<td>C1877_10655</td>
<td>hypothetical_protein</td>
<td>-1.65</td>
<td>1.43E-06</td>
</tr>
<tr>
<td>C1877_10605</td>
<td>sulfurtransferase_TusE</td>
<td>-1.65</td>
<td>3.30E-16</td>
</tr>
<tr>
<td>C1877_10670</td>
<td>dsrB:dissimilatory-type_sulfite_reductase</td>
<td>-1.66</td>
<td>1.58E-07</td>
</tr>
<tr>
<td>C1877_10685</td>
<td>glutamate-1-semialdehyde-aminomutase</td>
<td>-1.75</td>
<td>8.83E-16</td>
</tr>
<tr>
<td>C1877_10690</td>
<td>porphobilinogen_synthase</td>
<td>-1.78</td>
<td>4.43E-15</td>
</tr>
<tr>
<td>C1877_10665</td>
<td>putative_sulfate_exporter_family_transporter</td>
<td>-1.79</td>
<td>1.14E-09</td>
</tr>
<tr>
<td>C1877_10675:</td>
<td>dsrA:dissimilatory-type_sulfite_reductase_alpha_subunit</td>
<td>-1.87</td>
<td>9.90E-11</td>
</tr>
<tr>
<td>C1877_10660</td>
<td>hypothetical_protein</td>
<td>-1.92</td>
<td>1.63E-08</td>
</tr>
<tr>
<td>C1877_10695</td>
<td>uroporphyrinogen-III_C-methyltransferase</td>
<td>-1.95</td>
<td>2.68E-13</td>
</tr>
<tr>
<td>C1877_10650</td>
<td>hypothetical_protein</td>
<td>-1.97</td>
<td>1.99E-06</td>
</tr>
<tr>
<td>C1877_10700</td>
<td>hydroxymethylbilane_synthase</td>
<td>-1.98</td>
<td>1.92E-19</td>
</tr>
<tr>
<td>C1877_10680</td>
<td>4Fe-4S_dicluster_domain-containing_protein</td>
<td>-2.04</td>
<td>4.64E-09</td>
</tr>
</tbody>
</table>
4.2.8 *G. pamelaeae* likely harbors a distinct hydrocaffeic acid dehydroxylase

Our lysate assays revealed that *G. pamelaeae* 3C cells grown with DOPAC could also metabolize hydrocaffeic acid, suggesting the DOPAC dehydroxylase might have promiscuous activity (Figure 4.24). We also found that this organism metabolized hydrocaffeic acid in anaerobic cultures, despite not encoding for the Hcdh protein that we had biochemically verified in *E. lenta* A2 (Figure 4.27). We considered two possible explanations for this observation: either the hydrocaffeic acid dehydroxylase activity in *G. pamelaeae* 3C came from the DOPAC enzyme, or *G. pamelaeae* 3C encoded different dehydroxylase specific for hydrocaffeic acid. To distinguish between these possibilities, we performed another set of lysate assays, growing *G. pamelaeae* 3C with no substrate, DOPAC, or hydrocaffeic acid, and then exposing the lysates from these cultures to each of these substrates. Cells grown with no substrate did not metabolize any catechol. Consistent with previous assays, cells grown with DOPAC had promiscuous activity towards hydrocaffeic acid (Figure 4.33). However, cells grown with hydrocaffeic acid only dehydroxylated hydrocaffeic acid and did not turn over DOPAC (Figure 4.33). This pattern of metabolism was similar to *E. lenta* lysates from cells grown with hydrocaffeic acid as well as with the purified Hcdh protein (Figures 4.24 and 4.31), indicating that hydrocaffeic acid dehydroxylation is not performed by the same enzyme as DOPAC dehydroxylation in *Gordonibacter*. Thus, the promiscuity of cell lysates grown with DOPAC must derive from weak induction of an as-yet-unidentified hydrocaffeic acid dehydroxylase by DOPAC or from promiscuous biochemical activity of the DOPAC dehydroxylase. While these data suggest the existence of a hydrocaffeic acid dehydroxylase in *G. pamelaeae* 3C that differs from the biochemically confirmed *E. lenta* Hcdh, further RNA-sequencing experiments and protein purification are needed to identify the precise proteins involved.
Figure 4.33 Lysate assays suggest the presence of a *G. pamelaeae* 3C hydrocaffeic acid dehydroxylase distinct from *E. lenta* Hcdh.

*G. pamelaeae* 3C was grown anaerobically in BHI medium containing 10 mM formate. 0.5 mM of catechol was added to induce dehydroxylase expression, followed by anaerobic lysis and enzyme assays. Crude lysates were exposed to different substrates (500 µM) and reactions were allowed to proceed anaerobically for 20 hours. Assays mixtures were analyzed using LC-MS. Heat map represents the mean of three biological replicates (lysate reactions).

4.2.9 Norepinephrine and caffeic acid do not induce unique dehydroxylases

In our initial screen of gut Actinobacterial metabolism (Figures 4.17 and 4.18), we identified strains that metabolized caffeic acid and norepinephrine. To understand the molecular basis of these reactions, we performed additional RNA-sequencing experiments, exposing *E. lenta* A2 to caffeic acid and *E. lenta* 28b to DL-norepinephrine. Cultures of *E. lenta* A2 incubated with caffeic acid produce m-hppa as the end product, indicating an initial double bond reduction followed by dehydroxylation of hydrocaffeic acid (14). In this strain, caffeic acid highly upregulated (>100-fold) two genes annotated as fumarate reductases, which could potentially be involved in the reduction of the double bond to produce hydrocaffeic acid (Table 4.12). This substrate also induced low levels of expression of the gene encoding for Hcdh (<5-fold) and did
not upregulate any other molybdenum-dependent enzyme. This indicates some promiscuity in the transcriptional regulation of hcdh. Alternatively, hydrocaffeic generated from caffeic acid activates hcdh transcription in this setup.

Similarly, DL-norepinephrine induced Dadh but no other molybdenum-dependent enzyme in E. lenta 28b (Table 4.13). This is consistent with our observations that DL-norepinephrine upregulates dadh E. lenta A2 (Table 4.3) and suggests that Dadh likely is the main enzyme responsible for norepinephrine dehydroxylation among gut Actinobacteria. Interestingly, the same fumarate reductase that was the highest upregulated gene in response to DL-norepinephrine in E. lenta 28b shared 99.5% amino acid identity with the highest upregulated gene (also a fumarate reductase) in response to DL-norepinephrine in E. lenta A2 (Tables 4.3 and 4.13). This gene was not upregulated in response to any other substrate used in our study, suggesting a potentially specific in norepinephrine metabolism. One possibility could be benzyl alcohol reduction of norepinephrine to form dopamine, which could be directly dehydroxylated by Dadh. We are unaware of biochemically characterized microbial benzyl alcohol reductases, but an E. lenta enzyme performing the analogous reduction of a benzyl ether (also annotated as a fumarate reductase) was recently been described (19). Growing norepinephrine-dehydroxylation strains (as assessed by catechol dehydroxylation) with norepinephrine and assessing production of dopamine or m-tyramine is necessary step to explore this possibility. Direct purification or heterologous expression of this fumarate reductase enzyme represents another avenue for exploring its potential interactions with norepinephrine.
Table 4.12 Genes differentially expressed upon exposure of *E. lenta* A2 to caffeic acid.
The data represent the genes that are differentially expressed in response to exposure of *E. lenta* A2 to 0.5 mM caffeic acid relative to vehicle (>\(2\)-fold difference, FDR<0.1). The catalytic subunit of the putative hydrocaffeic acid dehydroxylase (*hcdh*) is highlighted in red. Data are from n=3 bacterial cultures for each condition.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta-A2_00699</td>
<td>fumarate reductase/succinate dehydrogenase flavoprotein</td>
<td>7.908394585</td>
<td>0</td>
</tr>
<tr>
<td>Elenta-A2_02359</td>
<td>fumarate reductase/succinate dehydrogenase flavoprotein</td>
<td>7.772412121</td>
<td>0</td>
</tr>
<tr>
<td>Elenta-A2_00830</td>
<td>fumarate reductase/succinate dehydrogenase flavoprotein</td>
<td>4.5200195</td>
<td>2.54E-207</td>
</tr>
<tr>
<td>Elenta-A2_02815</td>
<td><strong>Molybdopterin dinucleotide binding region (<em>hcdh</em>)</strong></td>
<td>2.321198159</td>
<td>8.78E-45</td>
</tr>
<tr>
<td>Elenta-A2_00336</td>
<td>fumarate reductase/succinate dehydrogenase flavoprotein</td>
<td>2.14718718</td>
<td>3.23E-38</td>
</tr>
<tr>
<td>Elenta-A2_01613</td>
<td>DL-carnitine dehydratase/bile acid inducible protein F</td>
<td>2.096939444</td>
<td>8.21E-87</td>
</tr>
<tr>
<td>Elenta-A2_01615</td>
<td>Enoyl-CoA_hydratase/isomerase</td>
<td>1.976136247</td>
<td>6.36E-51</td>
</tr>
<tr>
<td>Elenta-A2_01625</td>
<td>NADH:flavin_oxidoreductase/NADH_oxidase</td>
<td>1.88788179</td>
<td>6.98E-38</td>
</tr>
<tr>
<td>Elenta-A2_01624</td>
<td>acyl-CoA_dehydrogenase_domain_protein</td>
<td>1.882580004</td>
<td>8.95E-83</td>
</tr>
<tr>
<td>Elenta-A2_02358</td>
<td>transcriptional regulator, LuxR family</td>
<td>1.863077168</td>
<td>6.98E-38</td>
</tr>
<tr>
<td>Elenta-A2_02820</td>
<td>hypothetical_protein</td>
<td>1.82348334</td>
<td>1.82E-27</td>
</tr>
<tr>
<td>Elenta-A2_02817</td>
<td>hypothetical_protein</td>
<td>1.642194357</td>
<td>9.99E-23</td>
</tr>
<tr>
<td>Elenta-A2_02819</td>
<td>Cytoplasmic chaperone TorD family protein</td>
<td>1.574530363</td>
<td>1.17E-20</td>
</tr>
<tr>
<td>Elenta-A2_01608</td>
<td>transcriptional regulator, LyrR family</td>
<td>1.533686478</td>
<td>1.94E-29</td>
</tr>
<tr>
<td>Elenta-A2_01609</td>
<td>hypothetical_protein</td>
<td>1.456578673</td>
<td>4.02E-18</td>
</tr>
<tr>
<td>Elenta-A2_01619</td>
<td>hypothetical_protein</td>
<td>1.35869326</td>
<td>1.13E-15</td>
</tr>
<tr>
<td>Elenta-A2_02816</td>
<td>hypothetical_protein</td>
<td>1.217806749</td>
<td>1.40E-13</td>
</tr>
<tr>
<td>Elenta-A2_01618</td>
<td>NUDIX hydrolase</td>
<td>1.214364388</td>
<td>3.74E-13</td>
</tr>
<tr>
<td>Elenta-A2_02822</td>
<td>transcriptional regulator, LuxR family</td>
<td>1.209136901</td>
<td>1.85E-12</td>
</tr>
<tr>
<td>Elenta-A2_01480</td>
<td>2,5-didehydrogluconate reductase</td>
<td>-1.154925262</td>
<td>2.30E-21</td>
</tr>
</tbody>
</table>
Table 4.13 Genes differentially expressed upon exposure of *E. lenta* 28b to norepinephrine.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Annotation</th>
<th>log2FoldChange</th>
<th>p-value (FDR&lt;0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elenta_28b_02794</td>
<td>fumarate_reductase/succinate_dehydrogenase</td>
<td>4.00845705</td>
<td>3.12E-36</td>
</tr>
<tr>
<td>Elenta_28b_02709</td>
<td>molydopterin_dinucleotide-binding_region (<em>dadh</em>)</td>
<td>1.78431723</td>
<td>NA</td>
</tr>
<tr>
<td>Elenta_28b_00300</td>
<td>Amino acid adenylation domain protein</td>
<td>1.32448259</td>
<td>NA</td>
</tr>
<tr>
<td>Elenta_28b_01945</td>
<td>hypothetical protein</td>
<td>1.13821866</td>
<td>0.00012238</td>
</tr>
<tr>
<td>Elenta_28b_00095</td>
<td>hypothetical protein</td>
<td>1.12450135</td>
<td>1.34E-06</td>
</tr>
<tr>
<td>Elenta_28b_01308</td>
<td>membrane bound O-acyl transferase MBOAT family protein</td>
<td>1.02993366</td>
<td>NA</td>
</tr>
<tr>
<td>Eenta_28b_02110</td>
<td>Coproporphyrinogen III oxidase-like Fe-S oxidoreductase</td>
<td>1.01091969</td>
<td>1.89E-06</td>
</tr>
</tbody>
</table>
4.2.10 Model for *Eggerthella* and *Gordonibacter* dehydroxylases

Despite being expected to perform the same type of chemical reaction, the putative catechol dehydroxylases from *E. lenta* and *G. pamelaeae* that we identified through RNA-sequencing differ in sequence identity, genomic context, and predicted subunit composition (Figure 4.34 and Table 4.14). The dopamine, catechin and hydrocaffeic acid dehydroxylase genes from *E. lenta* (*dadh*, *cadh* and *hcdh*, respectively) likely encode for the catalytic subunits of membrane-bound protein complexes, as they co-localize with genes predicted to encode an electron shuttling 4Fe-4S ferredoxin and a putative membrane anchor (Figure 4.34) (52, 54). The molybdenum-dependent catalytic subunits of these enzymes are also all predicted to carry a Twin-Arginine-Translocation (TAT), suggesting they are trafficked to the membrane where the signal sequence is cleaved off (55). We found no peptide coverage of the TAT signal sequence in the proteomics experiment that identified *E. lenta* Hcdh, further confirming that this sequence is cleaved in the mature protein as in other membrane-anchored bis-MGD enzymes (Figure 4.30) (56). Though we only observed the catalytic subunits in our activity-based native purifications of Dadh and Hcdh from *E. lenta* A2, we speculate that sonication (the method of lysis) disrupted the associated between subunits during the preparation. We therefore propose that the enzymes exist as a multi-subunit complex inside the cell.

In contrast to the dehydroxylases encoded by *E. lenta* strains, *dodh* and similar genes from *G. pamelaeae* do not harbor a TAT signal sequence, encode for proteins that are smaller than the *E. lenta* enzymes, and co-localize with a gene predicted to encode a small electron shuttling 4Fe-4S protein, suggesting they are likely part of cytoplasmic protein complexes (Figure 4.34). These putative *G. pamelaeae* dehydroxylases are also encoded adjacent to members of the Major Facilitator Superfamily, transporters that may import or export the catechol substrates or
dehydroxylated metabolites (Figure 4.34) (37). Altogether, these data indicate the potential existence of distinct subtypes of molybdenum-dependent catechol dehydroxylases in gut Actinobacteria. The subunit composition, reactivity and maturation of bis-MGD enzymes will be discussed in further detail in Chapter 5.

Table 4.14 Accession numbers (Uniprot and Genbank) of putative *Eggerthella* and *Gordonibacter* dehydroxylases identified in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviated name</th>
<th>Organism</th>
<th>UNIPROT ID</th>
<th>Genbank protein ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dopamine dehydroxylase</td>
<td>dadh</td>
<td><em>E. lenta</em> A2</td>
<td>A0A369NIV7</td>
<td>RDC23575.1</td>
</tr>
<tr>
<td>Hydrocaffeic acid dehydroxylase</td>
<td>hcdh</td>
<td><em>E. lenta</em> A2</td>
<td>A0A369MIX7</td>
<td>RDC18391.1</td>
</tr>
<tr>
<td>Catechin dehydroxylase</td>
<td>cadh</td>
<td><em>E. lenta</em> A2</td>
<td>C8WLG2</td>
<td>RDC23615.1</td>
</tr>
<tr>
<td>DOPAC dehydroxylase</td>
<td>dodh</td>
<td><em>G. pamelae</em> 3C</td>
<td>A0A369LV65</td>
<td>RDB62136.1</td>
</tr>
<tr>
<td>Catechol lignan dehydroxylase</td>
<td>eldh</td>
<td><em>G. pamelae</em> 3C</td>
<td>A0A369M2I8</td>
<td>RDB65137.1</td>
</tr>
</tbody>
</table>
Figure 4.34 *Gordonibacter* and *Eggerthella* dehydroxylases display differences in genomic contexts and predicted subunit compositions.

A) Genomic contexts of putative catechol dehydroxylases. Opaque ORFs encode core genes (ranging from a to c). The scales are the same for the size of the *Eggerthella* and *Gordonibacter* genes.

B) Predicted subunit composition

C) Predicted subunit composition
Figure 4.34 (Continued).
Transparent ORFs encode genes not expected to directly participate in dehydroxylation. The *E. lenta* enzymes (Dadh, Hcdh, Cadh) are predicted to share similar subunit architectures. The molybdenum-dependent catalytic subunit (a, red) likely associates with an electron shuttling 4Fe-4S ferredoxin (b, purple) and a membrane anchor (c, beige). In contrast, the *G. pamelaeeae* predicted molybdenum-dependent enzymes are likely composed of a catalytic molybdenum subunit (a, red) and a small associated ferredoxin (b, purple). B) Model for *E. lenta* dehydroxylases. The gene encoding the predicted catalytic molybdenum-containing subunit (a, red) is co-localized with genes encoding a predicted 4Fe-4S cluster-containing protein (b, purple) and a predicted membrane anchor (c, gold) in the genome. The molybdenum-containing catalytic subunits is also predicted to contain an N-terminal 4Fe-4S cluster. Interestingly, the molybdenum-containing catalytic subunit also carries a Twin-Arginine-Translocation (TAT) signal sequence, suggesting it is trafficked to the cell membrane. There, it could form a membrane-anchored complex that receives electrons from quinone oxidation in the membrane. Electrons from quinone oxidation could move through the iron sulfur clusters of the various subunits to ultimately reduce the molybdenum cofactor that performs the 2e- reduction of the catechol substrate. C) Model for *Gordonibacter* dehydroxylases. The gene encoding the predicted catalytic molybdenum-containing subunit (a, red) is co-localized with a gene encoding a 4Fe-4S cluster-containing protein (b, purple) in the genome. These enzymes are not predicted to be translocated across the periplasm as they lack the TAT signal sequence. The molybdenum cofactor-containing subunit is not predicted to contain a canonical 4Fe-4S cluster. In contrast to the *E. lenta* dehydroxylases, the *Gordonibacter* dehydroxylases are predicted to be free-standing as there is no membrane anchor encoded nearby. Electrons from could derive from an as-yet-unknown donor and be transferred to the iron sulfur cluster of the predicted 4Fe-4S cluster-containing protein to reduce the molybdenum cofactor that performs the 2e- reduction of the catechol substrate.
4.2.11 Catechol dehydroxylases are variably distributed in metagenomes and correlate with metabolism by complex gut microbiota samples ex vivo

Our finding that catechol dehydroxylases and their associated metabolic activities are variably distributed among closely related gut Actinobacteria made us wonder whether human gut microbial communities would harbor similar genetic and metabolic diversity. To address this, Dr. Jordan Bisanz (Turnbaugh lab, UCSF) first searched >1800 publicly available human gut metagenomes (58) for dadh, hcdh, cadh, dodh, and the recently identified cldh (19) genes. Although found at generally low abundances, these catechol dehydroxylases were widely but variably distributed across these metagenomes. Dadh and hcdh were the most prevalent (in >70% and >90% of individuals, respectively), followed by cadh (30%), dodh (20%), and cldh (25%) (Figure 4.35). Notably, the prevalence of the different genes in metagenomes is consistent with their distribution among individual human gut Actinobacterial isolates in our strain library (Figure 4.27).
Figure 4.35 Catechol dehydroxylases display variable prevalence and abundance in human gut metagenomes. Prevalence estimates (%) as a function of minimum abundance reveals that the candidate dehydroxylases are variably distributed but prevalent in human gut metagenome samples at low relative abundances. This analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).
To assess the presence of catechol dehydroxylation in complex gut microorganisms, we incubated fecal samples from unrelated humans (n=12) *ex vivo* with hydrocaffeic acid, (+)-catechin, and deuterium-labeled dopamine and DOPAC and analyzed dehydroxylation by LC-MS/MS. In this experiment, we observed dehydroxylation of dopamine, hydrocaffeic acid, and DOPAC across the majority of subjects, indicating that metabolic activities of these low-abundance gut Actinobacteria are indeed prevalent (Figure 4.36). However, catechol metabolism varied between compounds and subjects, with some individuals metabolizing all compounds and some metabolizing none (Figure 4.36). (+)-Catechin was depleted without production of the corresponding dehydroxylated metabolites (data not shown). Although we did not identify the catechin metabolites in this experiment, this finding is consistent with catechin producing diverse metabolites when incubated with complex gut microbiota communities (59-61). It is possible that (+)-catechin is metabolized by pathways orthogonal to the benzyl ether reduction and subsequent dehydroxylation we have described in this chapter. However, it is also possible that metabolism involves benzyl ether reduction and reduction prior to processing by other pathways. Use of labeled (+)-catechin is necessary to understand the fate of this molecule in future *ex vivo* incubations and determine the importance and prevalence of the pathway for (+)-catechin metabolism that we have focused on in our work.
Human gut microbiota communities display variable metabolism of catechols ex vivo.

A-C) Metabolism of $d_1$-dopamine (A), hydrocaffeic acid (B), and $d_3$-DOPAC (C) by 12 unrelated human gut microbiota samples ex vivo. Samples were cultured anaerobically in BHI medium with 500 µM substrate for 72 hours and metabolism was analyzed by LC-MS/MS. Bars are mean % dehydroxylation ± SEM (n = 3 independent cultures for each fecal sample). Metabolizing *E. lenta* A2 or *G. pamelaeae* 3C strains were included as positive controls.
To investigate whether metabolic variability correlated with the presence of specific dehydroxylase enzymes, we further investigated DOPAC metabolism. We separated the 12 samples into a group of 9 metabolizers and 3 non-metabolizers (in which no biological replicate displayed dehydroxylation activity). qPCR enumeration in these cultures revealed that the abundance of the candidate DOPAC dehydroxylase gene dodh discriminated metabolizing and nonmetabolizing subjects (p<0.001, unpaired t-test) (Figure 4.37), and correlated significantly with dehydroxylation activity within the 9 metabolizers (Pearson’s correlation, r=0.73, R²=0.53, p<0.05) (Figure 4.37). Altogether, these data are consistent with our previous finding that dadh SNP status correlates with dopamine metabolism in human gut microbiotas ex vivo (see Chapter 3) and suggest that the candidate dehydroxylases may be active in complex gut communities.

**Figure 4.37 dodh abundance correlates with DOPAC metabolism by complex human gut communities ex vivo.**
A) The abundance of dodh correlates with DOPAC dehydroxylation in human gut microbiota samples. Data represent the average dodh abundance (as assessed with qPCR) across the three replicates for samples in Figure 4.38, Panel C. Results are mean abundance ± SEM (***P < 0.001, two-tailed t-test). B) The abundance of dodh among metabolizers correlates with DOPAC dehydroxylation. The data plotted represent the average dodh abundance (as assessed with qPCR) and the average % dehydroxylation across the three replicates for metabolizing samples in Figure 4.38, Panel C (all samples except 6,7, and 12). The line represents the best-fit trendline for linear regression. There was a significant linear correlation between dodh abundance and % DOPAC dehydroxylation (Pearson’s correlation, r=0.73, R²=0.53, p<0.05).
4.2.12 Catechol dehydroxylases are distinct from other molybdenum-dependent enzymes

We next investigated the relationship of catechol dehydroxylases to other characterized molybdenum-dependent enzymes. These enzymes bear no sequence homology to the only other biochemically characterized reductive aromatic dehydroxylase, the molybdenum-dependent enzyme 4-Hydroxybenzoyl CoA reductase (4-HCBR) that we first described in Chapter 3. Whereas 4-HCBR belongs to the xanthine oxidase family of molybdenum-dependent enzymes, the catechol dehydroxylases belong to the bis-MGD family of molybdenum-dependent enzymes, suggesting independent evolutionary origins (41, 50, 51). We will discuss the diversity, reactivity, and assembly of molybdenum-dependent enzymes in detail in Chapter 5.

To situate the newly discovered catechol dehydroxylases within the broader family of bis-MGD enzymes, we mapped our sequences onto a phylogenetic tree that we had adapted from previous work by Nitschke and co-workers (62). This phylogenetic analysis revealed that catechol dehydroxylases form a unique clade within the bis-MGD enzyme family, clustering away from pyrogallol hydroxytransferase (Pht), the only other bis-MGD enzyme known to modify the aromatic ring of a substrate (Table 4.15 and Figure 4.38) (63). The catechol dehydroxylases are instead most closely related to acetylene hydratase, an enzyme that adds water molecule to acetylene to generate acetaldehyde, providing a carbon source for the marine Proteobacterium Pelobacter acetylenicus (Figure 4.38) (51, 53, 62). A sequence similarity network (SSN) analysis using sequences of bis-MGD enzymes revealed distinct clusters of catechol dehydroxylases, further suggesting these enzymes are functionally different from known family members (Figure 4.39). The clustering of the dehydroxylases in the SSN did not simply reflect the phylogeny of the organisms because additional sequences from both Eggerthella and Gordonibacter were found in clusters containing distinct, biochemically characterized enzymes (Figure 4.40). In addition, we
found that the two catechol dehydroxylase-containing clusters also harbored sequences from organisms other than *Eggerthella* and *Gordonibacter* (Figure 4.42). Based on these data, we propose that catechol dehydroxylases are a distinct group of molybdenum-dependent enzymes.
Table 4.15 Accession numbers of bis-MGD enzymes used to generate the phylogenetic tree of the bis-MGD enzyme family (Figure 4.38).

<table>
<thead>
<tr>
<th>Uniprot accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>YP_361367</td>
</tr>
<tr>
<td>YP_387178</td>
</tr>
<tr>
<td>YP_435153</td>
</tr>
<tr>
<td>YP_583300</td>
</tr>
<tr>
<td>YP_638431</td>
</tr>
<tr>
<td>YP_682694</td>
</tr>
<tr>
<td>YP_741477</td>
</tr>
<tr>
<td>YP_743256</td>
</tr>
<tr>
<td>YP_840970</td>
</tr>
<tr>
<td>YP_916609</td>
</tr>
<tr>
<td>YP_920808</td>
</tr>
<tr>
<td>YP_931484</td>
</tr>
<tr>
<td>YP_964317</td>
</tr>
<tr>
<td>YP_001002743</td>
</tr>
<tr>
<td>YP_001013239</td>
</tr>
<tr>
<td>YP_001020903</td>
</tr>
<tr>
<td>YP_001055297</td>
</tr>
<tr>
<td>YP_001056256</td>
</tr>
<tr>
<td>YP_001056789</td>
</tr>
<tr>
<td>YP_001152746</td>
</tr>
<tr>
<td>YP_001153186</td>
</tr>
<tr>
<td>YP_001157040</td>
</tr>
<tr>
<td>YP_001233491</td>
</tr>
<tr>
<td>YP_001408699</td>
</tr>
<tr>
<td>YP_001409193</td>
</tr>
<tr>
<td>YP_001562139</td>
</tr>
<tr>
<td>YP_001585636</td>
</tr>
<tr>
<td>YP_001685488</td>
</tr>
<tr>
<td>YP_001951391</td>
</tr>
<tr>
<td>YP_001951406</td>
</tr>
<tr>
<td>YP_002016790</td>
</tr>
</tbody>
</table>
Table 4.1 (Continued).

<table>
<thead>
<tr>
<th>Uniprot accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>YP_002129528</td>
</tr>
<tr>
<td>YP_002457721</td>
</tr>
<tr>
<td>NP_147849</td>
</tr>
<tr>
<td>NP_902213</td>
</tr>
<tr>
<td>NP_906381</td>
</tr>
<tr>
<td>NP_906934</td>
</tr>
<tr>
<td>NP_906980</td>
</tr>
<tr>
<td>NP_907333</td>
</tr>
<tr>
<td>NP_907591</td>
</tr>
<tr>
<td>YP_310854</td>
</tr>
<tr>
<td>YP_466957</td>
</tr>
<tr>
<td>YP_524035</td>
</tr>
<tr>
<td>YP_524325</td>
</tr>
<tr>
<td>YP_571843</td>
</tr>
<tr>
<td>YP_871247</td>
</tr>
<tr>
<td>YP_001318866</td>
</tr>
<tr>
<td>YP_001319191</td>
</tr>
<tr>
<td>YP_001800219</td>
</tr>
<tr>
<td>YP_001942454</td>
</tr>
<tr>
<td>YP_002734116</td>
</tr>
<tr>
<td>YP_002796487</td>
</tr>
<tr>
<td>YP_002828931</td>
</tr>
<tr>
<td>YP_002882810</td>
</tr>
<tr>
<td>YP_002949599</td>
</tr>
<tr>
<td>YP_003099832</td>
</tr>
<tr>
<td>YP_003151500</td>
</tr>
<tr>
<td>YP_003161194</td>
</tr>
<tr>
<td>YP_003182037</td>
</tr>
<tr>
<td>YP_003303433</td>
</tr>
<tr>
<td>YP_003303806</td>
</tr>
<tr>
<td>YP_003311189</td>
</tr>
<tr>
<td>YP_003381996</td>
</tr>
<tr>
<td>Uniprot accession #</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>YP_003393294</td>
</tr>
<tr>
<td>YP_003863182</td>
</tr>
<tr>
<td>YP_001099944</td>
</tr>
<tr>
<td>YP_001423003</td>
</tr>
<tr>
<td>YP_001632440</td>
</tr>
<tr>
<td>YP_003699297</td>
</tr>
<tr>
<td>YP_001753784</td>
</tr>
<tr>
<td>A0A0F7JXH0_9GAMM</td>
</tr>
<tr>
<td>CAM58792</td>
</tr>
<tr>
<td>A0A0SCQF4_9BACT</td>
</tr>
<tr>
<td>A0A160FRG8_9BURK</td>
</tr>
<tr>
<td>AAU11841</td>
</tr>
<tr>
<td>AAQ19491</td>
</tr>
<tr>
<td>ABP63660</td>
</tr>
<tr>
<td>NP_213709</td>
</tr>
<tr>
<td>NP_951834</td>
</tr>
<tr>
<td>YP_001634827</td>
</tr>
<tr>
<td>NP_415742</td>
</tr>
<tr>
<td>NP_415991</td>
</tr>
<tr>
<td>EAT99379</td>
</tr>
<tr>
<td>A0A1C5V9P8_9FIRM</td>
</tr>
<tr>
<td>A0A1E4LUP6_9BURK</td>
</tr>
<tr>
<td>A0A1F2WFM2_9ACTN</td>
</tr>
<tr>
<td>A0A1F8PG51_9CHLR</td>
</tr>
<tr>
<td>A0A1M3JW53_9BURK</td>
</tr>
<tr>
<td>A0A1M6E5L2_9FIRM</td>
</tr>
<tr>
<td>A0A1M6FDE7_9FIRM</td>
</tr>
<tr>
<td>A0A1L3GSN8_9DELT</td>
</tr>
<tr>
<td>A0A1W6CT59_9PROT</td>
</tr>
<tr>
<td>YP_429324</td>
</tr>
<tr>
<td>AAR05656</td>
</tr>
<tr>
<td>NP_460720</td>
</tr>
</tbody>
</table>
Table 4.15 (Continued).

<table>
<thead>
<tr>
<th>Uniprot accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0A1G6AWF4_EUBOX</td>
</tr>
<tr>
<td>A0A1Z8A857_9GAMM</td>
</tr>
<tr>
<td>AIOA_ALCFA</td>
</tr>
<tr>
<td>AIOA_HERAR</td>
</tr>
<tr>
<td>PGTL_PELAC</td>
</tr>
<tr>
<td>A0A2A4UZ40_9GAMM</td>
</tr>
<tr>
<td>A0A369NIV7</td>
</tr>
<tr>
<td>A0A369MIX7</td>
</tr>
<tr>
<td>C8WLG2</td>
</tr>
<tr>
<td>A0A369LV65</td>
</tr>
<tr>
<td>A0A369M2I8</td>
</tr>
<tr>
<td>A0A369M6P4</td>
</tr>
<tr>
<td>A0A369M3A8</td>
</tr>
</tbody>
</table>
Figure 4.38 Catechol dehydroxylases are distinct from other bis-MGD enzymes.
Phylogenetic analysis of newly discovered catechol dehydroxylases reveals a unique evolutionary origin and relationship to acetylene hydratase. Psr = polysulfide reductase; Arr = arsenate reductase; Fdh = formate dehydrogenase; Pht = phloroglucinol transhydroxylase; Dms = DMSO reductase; Tmr = TMAO reductase; Aro = arsenite oxidase; Ebd = ethylbenzene dehydrogenase; Pcr = perchlorate reductase; Nar = nitrate reductase; Cdh = catechol dehydroxylase. The maximum likelihood tree was constructed using sequences from (62) as well as additional family members and reproduced the previously reported phylogeny of this enzyme family. Black circles on branches indicate bootstrap values greater than 0.7.
Figure 4.39 Sequence similarity network of the bis-MGD enzyme family reveals that the *Gordonibacter* and *Eggerthella* dehydroxylases belong to distinct, uncharacterized clusters. The SSN was constructed using the bis-MGD binding domain enzyme superfamily (PF01568) from Uniprot. Nodes represent proteins with 75% amino acid sequence identity. SSN displayed with an e-value threshold of $10^{-167}$. Clusters containing the representative biochemically characterized enzymes included in the phylogenetic analysis presented in Figure 4.40 are indicated with black boxes (a cluster includes all connected nodes). Clusters with blue nodes contain the dehydroxylases identified in this study. Red nodes represent sequences that have been reviewed in the Uniprot database and have been annotated with evidence for their function at the genetic, transcriptional, or biochemical level.
Figure 4.40 Sequence similarity network of the bis-MGD enzyme family reveals that *Gordonibacter* and *Eggerthella* bis-MGD enzymes are widely distributed across the enzyme family.

The SSN is the same as the one presented in Figure 4.39. Clusters containing the biochemically characterized enzymes included in the phylogenetic analysis presented in Figure 4.40 as well as the candidate dehydroxylases in our study are indicated with black boxes (a cluster includes all connected nodes). Red nodes represent sequences from *Eggerthella*, while blue nodes represent sequences from *Gordonibacter*. This network indicates that the clustering does not simply reflect the phylogeny of the host organisms as *Gordonibacter* and *Eggerthella* enzymes are widely distributed across the network.
4.2.13 GNN analysis of *Eggerthella* and *Gordonibacter* dehydroxylases

We next used genome neighborhood network (GNN) analysis, a tool that enables visualization and quantification of the broader genomic context the proteins contained within clusters of an SSN (64). We subjected the clusters harboring the *Gordonibacter* and *Eggerthella* dehydroxylases Cadh, Hcdh, Dadh, Dodh, and Cldh (clusters L and M, respectively in Figure 4.40) to GNN analysis. We used a co-occurrence of 20%, which specifies the threshold above which a specific gene family is considered to be associated with the gene of interest. We also used a neighborhood size of 10, which specifies that only genes within a 10 gene distance (and no further away) from the gene of interest would be considered “associated” genes.

Consistent with our previous analysis of the genomic contexts of newly discovered catechol dehydroxylases (Figure 4.34), *Gordonibacter* and *Eggerthella* dehydroxylases generally co-localized with distinct protein families (Figure 4.41). For example, *Gordonibacter* genes are frequently associated with MFS transporters, iron-sulfur cluster proteins, as well as other molybdenum-dependent enzymes. Other associated gene families included transcriptional regulators (HTH, sigma activator type), B$_{12}$-binding proteins, and uroporphyrinogen decarboxylase-type enzymes. In contrast, members of the SNN cluster harboring catechol dehydroxylases from *Eggerthella* most frequently associate with LuxR type transcriptional regulators (GerE type), nitrate reductase delta subunit type proteins (likely a dedicated chaperone installing the molybdopterin cofactor in the dehydroxylase enzyme), and iron-sulfur cluster proteins. Other gene families include LysR type transcriptional regulators (harboring an HTH domain) and DmsC-type proteins, which likely represents the membrane subunit of the dehydroxylase complex (Figure 4.42). Based on further visual inspection of the GNN output, we created models of an “average” operonic architecture for *Gordonibacter* and *Eggerthella*-type
dehydroxylases (Figure 4.42). This architecture is largely consistent with the one proposed based on our initial analysis of Dadh, Cadh, Hcdh, Dodh, and Cldh (Figure 4.34). The broader differences between *Gordonibacter* and *Eggerthella*-type dehydroxylases suggests these enzymes could have different functions, cellular localizations, and maturation pathways. The subunit composition, functions, and maturation of bis-MGD enzymes is discussed in detail in Chapter 5.
Figure 4.41 GNN analysis of *Gordonibacter* and *Eggerthella* type dehydroxylases.
We subjected the clusters from the bis-MGD enzyme family SSN harboring the *Gordonibacter* and *Eggerthella* dehydroxylases Cadh, Hcdh, Dadh, Dodh, and Cldh (clusters L and M, respectively in Figure 4.42) to GNN analysis, using a co-occurrence of 20 and a neighborhood size of 10.
Figure 4.41 (Continued).

A) Shows the gene families associated with the Gordonibacter dehydroxylases. The red diamond represents the SSN cluster, and the spokes coming off it represents the broader protein families frequently co-localized with this cluster. The size of these circles indicates the relative frequency with which the protein family is associated with the cluster of interest. MFS = MFS type transporter; Fer4 = iron-sulfur cluster proteins, Uro-D = Uroporphyrinogen decarboxylase-type proteins, Pterin represents pterin-binding sequences, Sigma54r = transcriptional activator, none = sequences with no annotation. B) Shows the gene families associated with the Eggerthella dehydroxylases. The blue diamond represents the SSN cluster, and the spokes coming off it represents the broader protein families frequently co-localized with this cluster. The size of these circles indicates the relative frequency with which the protein family is associated with the cluster of interest. Fer4 = electron-shuttling ferredoxin sequences, DmsC = membrane anchor subunit, nitrate reductase delta subunit = dedicated chaperone inserting the molybdopterin co-factor, GerE = luxR-type transcriptional regulator, LysR = a different type of transcriptional regulator, AAA = proteins containing an AAA domain (involved in a diversity of processes), and FAD binding = sequences predicted to bind FAD as a cofactor. none = sequences with no annotation.
Figure 4.42 GNN-predicted average operonic architecture of *Gordonibacter* and *Eggerthella* type dehydroxylases.
The GNN output was visually inspected, allowing us to create a model for the average operon for each dehydroxylase type based on the frequency and relative position of frequently co-localizing genes. A non-dashed, opaque open reading frame indicates a core gene in terms of its location and direction, found in the vast majority (approximately 90% of operons). A transparent shape indicates a sequence is often (>50% of cases) found in the vicinity of a putative catechol dehydroxylase, and the dash indicates that the location and direction of the gene can vary between operons. *Gordonibacter* sequences, which includes *cldh* and *dodh*, are quite diverse and usually consist of a freestanding molybdenum-dependent enzyme. This enzyme is usually associated with other bis-MGD enzymes, with membrane MFS transporters, and sometimes a small ferredoxin. However, the direction and proximity of these enzymes vary. *Eggerthella* sequences, which includes *dadh*, *hcdh*, and *cadh*, are comprised of a core three-gene operon containing the catalytic subunits of the bis-MGD enzyme and its associated electron transport and membrane subunits. A strong association with a transcriptional regulator is also observed, which could be influencing the specificity in transcription of these enzymes. There is often a putative bis-MGD insertion chaperone somewhere in the vicinity as well, but the precise location and direction vary.
4.2.14 Catechol dehydroxylases are widely distributed among sequenced microbes

To assess the diversity of putative catechol dehydroxylases, we queried the NCBI nucleotide database and our collection of Actinobacterial genomes for homologs of the *Eggerthella* and *Gordonibacter* enzymes. Phylogenetic analyses of the resulting sequences revealed a large diversity of potential catechol dehydroxylases, including numerous uncharacterized enzymes encoded in individual *Gordonibacter* and *Eggerthella* genomes (Figure 4.43). *Gordonibacter* harbored as many as 39 such enzymes, while *Eggerthella* strains harbored as many as 14 homologs (Figure 4.44). This analysis highlights that catechol dehydroxylases likely have diversified within these closely related gut Actinobacteria, that individual gut Actinobacterial strains can likely metabolize a range of different catechols, and that many substrate-enzyme pairs remain to be discovered.

This analysis also revealed that potential catechol dehydroxylases are not restricted to human-associated Actinobacteria and are instead part of a larger group of bis-MGD enzymes present in diverse bacteria and even Archaea (Figure 4.43). These organisms come from mammal-associated, plant-associated, soil, and aquatic habitats. Notable organisms encoding putative dehydroxylases include soil-dwelling Streptomycetes (65, 66), the industrially important anaerobe *Clostridium ljungdahlii* (67), and a large number of anaerobic bacterial genera known for their ability to degrade aromatic compounds, including *Azoarcus, Thauera, Desulfobacula, Geobacter, Desulfomonile*, and *Desulfotobacterium* (Figure 4.43) (68-76). These latter genera degrade aromatics through reactions such as dehalogenation and demethoxylation but have not been reported to dehydroxylate catechols. The presence of catechol dehydroxylase homologs in these strains suggests catechol dehydroxylation might be an unappreciated aspect of aromatic degradation by environmental bacteria. More broadly, the presence of similar enzymes in gut and
environmental microbes likely reflects the availability of catechol substrates in many different environments (Figure 4.43).

**Figure 4.43 Catechol dehydroxylases are widely distributed across sequenced microbes.**
Maximum likelihood phylogenetic tree for catechol dehydroxylase homologs identified by querying 26 gut Actinobacterial genomes and the NCBI nucleotide collection for Dadh and Cldh homologs (see methods for details).
Figure 4.43 (Continued).
The color of the lines indicates the phylogeny of the organism harboring the homolog. The color of the border indicates the primary habitat from which the organism was originally isolated. Numbers at the end of the branches indicate highlighted sequences, which are specified in the legend above. Unchar. = uncharacterized. The color of the organism name matches the phylogeny of the organism. DMSO reductase from *E. coli* (sequence #15) was used as an outgroup to root the tree. All of the sequences highlighted in the figure are mentioned in the main text. Black circles on branches indicate bootstrap values greater than 0.7.
Figure 4.44 Number of catechol dehydroxylase homologs among gut Actinobacterial isolates used in our study.

Cldh was used to perform a tBLASTn search of the *Gordonibacter* genomes. Sequences with an amino acid ID of >30% and e-value of e-34 were retrieved as potential hits. Dadh was used to perform a tBLASTn search of the *Eggerthella* genomes. Sequences with an amino acid ID of >30% and e-value of e-34 were retrieved as potential hits.
As the vast majority of catechol dehydroxylase homologs remain uncharacterized, it is difficult to assign the biochemical activities of the major clades and define the characteristic features of these enzymes. However, we are confident that at least some portion of the sequences captured in this analysis are true catechol dehydroxylases. First, we found that representative sequences from across our phylogenetic tree are more closely related to acetylene hydratase and the *Gordonibacter* and *Eggerthella* dehydroxylases than to any other member of the bis-MGD enzyme family, indicating shared evolutionary origins (Table 4.16 and Figures 4.45 and 4.46). This clustering also suggests some of the sequences captured in our search might be true acetylene hydratases. This idea is also supported by the presence of a group of *Gordonibacter* sequences in the sub-clade harboring the biochemically characterized acetylene hydratase in Figure 4.43. The presence of a potential acetylene hydratase (70% amino acid identity to acetylene hydratase) in human gut *Gordonibacter* is surprising given that acetylene hydratase is an ancient enzyme thought to have evolved on primordial earth when acetylene was more abundant in the atmosphere (77, 78). The abundance of acetylene in the human gut microbiota is unknown. Further structural and biochemical work is required to determine the precise functions of the dehydroxylase homologs captured in our analysis.
<table>
<thead>
<tr>
<th>Representative sequence</th>
<th>Organism</th>
<th>Uniprot accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>Gordonibacter pamelaeae</em></td>
<td>A0A369M3J7</td>
</tr>
<tr>
<td>2</td>
<td><em>Gordonibacter pamelaeae</em></td>
<td>A0A369M6U3</td>
</tr>
<tr>
<td>3</td>
<td><em>Desulfitobacterium metallireducens</em></td>
<td>W0E998</td>
</tr>
<tr>
<td>4</td>
<td><em>Gordonibacter sp.</em></td>
<td>A0A369LVG2</td>
</tr>
<tr>
<td>5</td>
<td><em>Clostridium ljungdahlii</em></td>
<td>D8GM28</td>
</tr>
<tr>
<td>6</td>
<td><em>Eggerthella sp.</em></td>
<td>F0HRV1</td>
</tr>
<tr>
<td>7</td>
<td><em>Eggerthella lenta</em></td>
<td>C8WM38</td>
</tr>
<tr>
<td>8</td>
<td><em>Eggerthella lenta</em></td>
<td>A0A369MIX7</td>
</tr>
<tr>
<td>9</td>
<td><em>Eggerthella lenta</em></td>
<td>A0A369NI47</td>
</tr>
<tr>
<td>10</td>
<td><em>Streptomyces zelensis</em></td>
<td>A0A1W6EUS9</td>
</tr>
<tr>
<td>11</td>
<td><em>Thauera aromatica</em></td>
<td>A0A088SRF0</td>
</tr>
<tr>
<td>12</td>
<td><em>Gordonibacter pamelaeae</em></td>
<td>A0A369LSI8</td>
</tr>
<tr>
<td>13</td>
<td><em>Gordonibacter pamelaeae</em></td>
<td>D6E7S9</td>
</tr>
<tr>
<td>14</td>
<td><em>Denitrobacterium detoxificans</em></td>
<td>A0A172RZZ3</td>
</tr>
<tr>
<td>15</td>
<td><em>Bradyrhizobium canariense</em></td>
<td>A0A1H1VHY4</td>
</tr>
<tr>
<td>16</td>
<td><em>Thiocapsa roseopersicina</em></td>
<td>X5FC28</td>
</tr>
<tr>
<td>17</td>
<td><em>Thioflavicoccus mobilis 8321</em></td>
<td>L0GWW7</td>
</tr>
<tr>
<td>18</td>
<td><em>Ferroglobus placidus</em></td>
<td>D3S147</td>
</tr>
</tbody>
</table>
Figure 4.45 Phylogenetic analysis of representative catechol dehydroxylase homologs highlighting sequences included in further analyses.
This phylogenetic tree is the same as in Figure 4.43. However, here the numbers indicate homologs that were selected as representative sequences across the tree for further phylogenetic analysis. The accession numbers of these sequences can be found in Table 4.16.
Phylogenetic analysis of putative catechol dehydroxylase sequences from diverse bacterial phyla reveals a relationship with the biochemically characterized molybdenum-dependent enzyme acetylene hydratase.

The representative sequences from Figure 4.45 were added to the tree containing biochemically characterized members of the bis-MGD enzyme family (Figure 4.40). Psr = polysulfide reductase; Arr = arsenate reductase; Fdh = formate dehydrogenase; Pht = phloroglucinol transhydroxylase; Dms = DMSO reductase; Tmr = TMAO reductase; Aro = arsenite oxidase; Ebd = ethylbenzene dehydrogenase; Pcr = perchlorate reductase; Nar = nitrate reductase. Black circles on branches indicate bootstrap values greater than 0.7. This analysis revealed that all of these representative catechol dehydroxylase homologs are more closely related to each other and the acetylene hydratase sequences than they are to any other biochemically characterized bis-MGD enzyme family member. Acetylene hydratases are shown in blue and the representative sequences selected from Figure 4.45 are shown in black.
In addition to our bioinformatic analysis, recent genetic studies have implicated several homologs from environmental bacteria in catechol dehydroxylation. For instance, a putative dehydroxylase is present in the *Streptomyces* biosynthetic gene clusters that produce the potent anti-tumor compounds yatakemycin and CC-1065 (Figures 4.45 and 4.49) (65, 66). This protein is predicted to have a partner electron transfer shuttling protein, similar to *E. lenta* and *Gordonibacter* enzymes (65). Gene knock-out and complementation studies revealed the dehydroxylase homolog is essential for CC-1065 production and likely catalyzes reductive dehydroxylation of a late-stage biosynthetic intermediate prior to installation of the DNA-alkylating cyclopropane motif (Figure 4.47) (66). The dehydroxylated catechol in CC-1065 originally comes from L-dopa, a starter unit that is incorporated early during biosynthesis (65). Interestingly, the chemical motif in CC-1065 that results from dehydroxylation is also present in other potent anti-tumor natural products, indicating this chemical strategy may be relevant beyond the biosynthesis of CC-1065 (65). Finally, the observation that catechol dehydroxylase homologs are present in additional Streptomyces genomes may suggest a broader role for dehydroxylation in natural product biosynthesis (Figure 4.43). Preliminary GNN analysis of *Streptomyces* dehydroxylase homologs (conducted by Dr. Chip Le in our laboratory) supports this idea, indicating that many of these genes are present in biosynthetic gene clusters (data not shown). For example, a homolog of unknown function (named *belN*, sharing 23 % amino acid identity over 86% query coverage with *dodh*) is present in the biosynthetic gene cluster that produces the proteasome inhibitor belactosin (79).
A putative catechol dehydroxylase is critical for the biosynthesis of the anti-tumor compound CC-1065. The original results were reported by Tang and co-workers (65).

**Figure 4.47** A putative catechol dehydroxylase is critical for the biosynthesis of the anti-tumor compound CC-1065.

The original results were reported by Tang and co-workers (65).
Another catechol dehydroxylase homolog is present in the 3,5-dihydroxybenzoic acid (3,5-DHB) degradation operon from the anaerobic soil Proteobacterium *Thaurea aromatica* (Figures 4.43 and 4.58). The proposed pathway for 3,5-DHB degradation includes two catecholic intermediates, which induce the expression of the candidate dehydroxylase gene (74, 76). Moreover, strains lacking this enzyme exhibit impaired growth on 3,5-DHB as a sole carbon source (74). While the biochemical function of this dehydroxylase homolog is still unknown, these data implicate the candidate dehydroxylase (named *T. aromatica* dehydroxylase, or Tardh) in potential metabolism of one of the two catecholic intermediates involved in this pathway. Based on knock-out experiments and genetics, Marques and co-workers recently proposed that this enzyme is involved in decarboxylation of 1,3,5-trihydroxybenzoic acid, although this claim remains unsubstantiated at the biochemical level (74, 76, 80). Further biochemical studies are necessary to verify the function of this gene. Based on these analyses, we conclude that the catechol dehydroxylases harbor vast uncharacterized diversity that contributes to both primary and secondary metabolic pathways in habitats beyond the human gut.

**Figure 4.48 A putative catechol dehydroxylase is important for *T. aromatica* AR-1 growth on 3,5-dihydroxybenzoic acid.**
The figure shows the proposed pathway for 3,5-dihydroxybenzoic acid (3,5-DHB) degradation. The original results were reported by Marques and co-workers (74). Cases where genes involved in the transformation are known are highlighted above the arrow.
4.2.15 Catechol dehydroxylase reactivity is present across the gut microbiotas of mammals representing distinct diets and phylogenetic origins

Our phylogenetic analysis suggested that catechol dehydroxylase activity is present in a range of microbial habitats, making us curious whether we could detect this metabolism in additional microbial communities. As a first step, we explored catechol dehydroxylation by gut microbiotas of non-human mammals. We assembled a panel of gut microbiota samples from 12 different mammals representing diverse phylogenetic origins and diets (3 individuals per mammal) (81) (Figures 4.49 and 4.50). These samples were generously provided by Dr. Aspen Reese, a Junior Fellow working with Dr. Rachel Carmody at Harvard University. We cultured these gut communities anaerobically ex vivo, assessed metabolism using a colorimetric assay, and confirmed potential hits using LC-MS/MS (Figure 4.49). We observed catechol dehydroxylation across the gut microbiotas of mammals spanning different diets and phylogenies (Figure 4.50). Hydrocaffeic acid dehydroxylation occurred in >50% of species, while dopamine and (+)-catechin metabolism were observed in 5/12 and 4/12 animals, respectively (Figure 4.50). DOPAC was only metabolized by the rat gut microbiota, which was the only community that had activity towards all compounds tested. While a larger sample size is required to reach clear conclusions about possible links between metabolism of specific catechols and individual mammal gut microbiotas, our results clearly demonstrate that catechol dehydroxylation is found in distantly related mammal gut microbiotas that have large differences in species composition and gene content (81-83). This finding further reinforces the relevance of catechol dehydroxylation to variety of different microbial habitats.
Figure 4.49 Screen for catechol dehydroxylation by gut microbiota samples from diverse mammals.
A) Dehydroxylation of dopamine, DOPAC, (+)-catechin, and hydrocaffeic acid by gut microbiota samples from mammals spanning distinct diets and phylogenetic groups.
Figure 4.49 (Continued).
Gut communities from \( n = 12 \) different mammals and \( n = 3 \) individuals per animal were cultured anaerobically for 96 hours in basal medium with 0.5 mM catechol at 37 °C. Metabolism was then assessed using a colorimetric assay for catechol detection. Data were normalized to the sterile control. Each dot represents a different individual for each animal, and the red color indicates samples that were selected for further LC-MS/MS analysis. Bars display the mean and standard deviation. B) \( m \)-Tyramine mass peak area in select mammalian gut microbiota samples incubated with dopamine and selected for LC-MS/MS analysis. Samples that displayed catechol depletion in A) performed dehydroxylation of dopamine. C) \( m \)-HPAA mass peak area in select mammalian gut microbiota samples incubated with DOPAC and selected for LC-MS/MS analysis. Only the rat microbiota (individual 1) had activity towards DOPAC. D) Dehydroxylated catechin derivative mass peak area in select mammalian gut microbiota samples incubated with (+)-catechin and selected for LC-MS/MS analysis. All samples selected for further analysis except the sheep microbiota displayed the full two-step conversion of (+)-catechin into the dehydroxylated derivative. The red asterisk indicates that the wolf microbiota had activity based on LC-MS/MS even though the mass peak area was lower than what is clearly visible with the current scale of the Y-axis. E) \( m \)-HPPA mass peak area in select mammalian gut microbiota samples incubated with hydrocaffeic acid and selected for LC-MS/MS analysis. The screen for metabolism of all compounds across the animals was performed once.
Gut microbiotas of mammals representing distinct diets and phylogenetic origins can dehydroxylate catechols.

Catechol dehydroxylation of dopamine, DOPAC, (+)-catechin, and hydrocaffeic acid by gut microbiota samples from mammals spanning distinct diets and phylogenetic groups. Gut communities from 12 different mammals and 3 individuals per animal were cultured anaerobically for 96 hours in basal medium with 0.5 mM catechol at 37 ºC. The results summarize animals and individuals where the known dehydroxylation pathways examined in human gut Actinobacteria took place, as assessed by LC-MS/MS. Red indicates that metabolism took place in at least one of the individuals, and black indicates lack of metabolism, as assessed by the detection of the dehydroxylated metabolite using LC-MS/MS. The experiment was performed once. The phylogenetic tree was created using the aptg plugin in R and missing branches were added manually based on mammalian phylogeny. The icons were adapted under a Creative Commons license (https://creativecommons.org/licenses/by/3.0/) at phylopic (http://phylopic.org), including Alpaca logo (made my Steven Traver), Bison (Lukasiniho), Cow (Steven Traver), Dog (Tracy A Heath), Fox (Anthony Caravaggi), Guinea pig (Zimices), Mouse (Madeleine Price Ball), Pig (Steven Traver), Rabbit (Steven Traver), Rat (Rebecca Groom), Sheep (Zimices), and Wolf (Tracy A Heath).
4.3 Discussion

For many decades the human gut microbiota has been known to dehydroxylate catechols, but the molecular basis of this enigmatic transformation has remained largely unknown. In this Chapter, we characterized the specificity and regulation of a gut bacterial enzyme that dehydroxylates dopamine (Dadh). We then used this knowledge to identify candidate enzymes that dehydroxylate additional host- and plant-derived small molecules. Together, the catechol dehydroxylases represent a previously unappreciated group of molybdenum-dependent enzymes that is present in diverse microbial phyla and environments. Our studies of Dadh activity and transcriptional regulation revealed a high specificity for dopamine and norepinephrine. While these two catecholamine neurotransmitters were metabolized by many strains in our collection of >25 gut Actinobacterial isolates, not a single strain clearly dehydroxylated epinephrine. These findings are notable because whereas dopamine and norepinephrine are biosynthesized by the host in the gastrointestinal tract (42, 84), epinephrine is not present in the gut due to the absence of the N-methyltransferase (85, 86). This observation supports the hypothesis that the physiological role of Dadh is to enable neurotransmitter metabolism by \textit{E. lenta} in the human gut (Figure 4.51). This idea is also consistent with recent observations of gut bacteria using specific host neurotransmitters for growth (87, 88).
**Figure 4.51 Host biosynthesis and gut Actinobacterial metabolism of catecholamines.**
Host enzymes present in the gastrointestinal tract are highlighted in blue, while enzymes thought to be absent in the gut are highlighted in red. TYH = tyrosine hydroxylase; AADC = aromatic amino acid decarboxylase; DBH = dopamine-beta hydroxylase; PNMT = phenylethanolamine N-methyltransferase. L-epinephrine is not biosynthesized in the gut due to the absence of detectable PNMT expression in gastrointestinal cells and tissues (86). The observation that gut Actinobacteria metabolize dopamine and L-norepinephrine suggests that dehydroxylation has evolved to degrade physiologically available neurotransmitters.
To our knowledge, Dadh represents the first catecholamine-metabolizing enzyme discovered from a human gut commensal. However, as mentioned in Chapter 3 of this thesis, interactions between catecholamines and intestinal pathogens are well-characterized and have long been known as key players in virulence and infection (89, 90). Whereas pathogenic organisms such as *Escherichia coli*, *Yersinia enterocolitica*, and *Salmonella enterica* require the intact catechol group of dopamine and norepinephrine to sequester iron and boost growth (89-92), we propose that *E. lenta* uses these molecules as electron acceptors. Thus, Dadh might represent a novel strategy by which gut bacteria take advantage of catecholamines present in the gastrointestinal tract (42, 84). Understanding the interplay between pathogenic and commensal interactions with catecholamines is an intriguing avenue for further research.

In addition to characterizing Dadh, we discovered candidate dehydroxylases that metabolize (+)-catechin, hydrocaffeic acid, and DOPAC. We also partially purified the hydrocaffeic acid dehydroxylase to confirm its involvement in this reaction. While further biochemical studies are important for validating the activities of the remaining enzymes, identification of candidate genes involved in dehydroxylation of host and diet-derived catechols significantly advances our understanding of how these molecules are processed in the gut microbiota. While many dietary molecules are degraded by the gut microbial community as a whole, our findings provide rare molecular information about these processes, affording researchers opportunities to elucidate biological links between gut microbial phenol metabolism and human health. Our current data support a working model in which catechol dehydroxylation is performed by distinct enzymes that are specialized for individual substrates, and we identified large numbers of uncharacterized dehydroxylases encoded within individual *Eggerthella* and *Gordonibacter* genomes (Figures 4.43 and 4.45). Notably, *Eggerthella* and *Gordonibacter* encode
for among the most molybdenum-dependent enzymes among any sequenced microbial isolates (54), and our analysis suggests many of these are likely catechol dehydroxylases (Figure 4.45), hinting at an expansion of this group of enzymes among human gut Actinobacteria.

While it remains to be seen whether uncharacterized dehydroxylase homologs are also specific for distinct substrates, this type of diversification of closely related enzymes indicates a potentially important role for catechol dehydroxylation in the human gut microbiota. Expansion of enzyme families within specific clades of gut microbes is well-characterized in the context of polysaccharide metabolism. For example, individual human gut Bacteroides strains isolates harbor hundreds of polysaccharide utilization loci but upregulate only a subset of genes in response to distinct substrates (93-96). This transcriptional regulation and biochemical specificity enable utilization of various host- or plant-derived carbon sources depending on their availability (97-99). The diversity of catechol dehydroxylases might have evolved for a similar purpose, providing a biochemical arsenal that enables Actinobacteria to use a range of different electron acceptors whose availability depends on the diet and/or physiology of the host.

In addition to identifying the molecular basis of a prominent microbial transformation, our study underscores how enzyme discovery can help to dissect the metabolic diversity of gut microbial strains and communities. Although previous studies had linked certain dehydroxylation reactions to individual gut Actinobacteria (18, 19, 39), we have found that specific catechol dehydroxylases are variably distributed among closely related strains and human gut metagenomes. These findings reinforce the idea that gut microbial phylogeny is often not predictive of functional capabilities (44, 49, 100, 101). Additionally, we noticed that the prevalence of the different dehydroxylation reactions among human and animal gut microbiota samples reflected their distribution among individual Actinobacterial strains, with hydrocaffeic
acid metabolism being the most prevalent across all strains and species, and DOPAC dehydroxylation being the least prevalent. This may suggest that the strain-level variability in dehydroxylases is important for metabolism both within humans and other mammalian species.

While the evolutionary forces shaping the distribution of specific dehydroxylases within gut bacterial strains and complex gut communities remain unknown, our study provides a starting point for future investigations of catechol metabolism and its biological consequences in the human gut microbiota and beyond.

Our observations that dehydroxylases are present in known metabolic pathways in environmental organisms implicate these enzymes in both primary and secondary metabolism beyond the human gut microbiota. However, a future challenge will be to identify the substrates and ecological functions of unknown dehydroxylase homologs encoded by both gut and environmental strains. In our colorimetric screen described in this Chapter, we evaluated dehydroxylation of 20 substrates by human gut Actinobacteria. However, it is possible that the substrate scope of unknown dehydroxylases extends far beyond these commercially available compounds, explaining why we only assigned a putative function to a small fraction of sequences in the present study. For example, some substrate may derive from methoxylated compounds, which are prominent in plants and are also produced by humans (19, 102) (Figure 4.55). These compounds would require microbial demethylation to unmask a free catechol prior to dehydroxylation, a process that is observed in the inter-species, multi-step metabolism of dietary lignans (19). A more detailed discussion of potential substrates will be provided in Chapter 6 of this thesis. Identifying the substrates of uncharacterized catechol dehydroxylases, as well as how these genes are regulated, could shed light on how gut organisms have adapted to molecules
produced and ingested by the host and uncover the biochemical functions and ecological roles of dehydroxylation among environmental microbes.
4.4 Experimental methods

4.4.1 General materials and methods

The following chemicals were used in this study: tetracycline (Sigma Aldrich, catalog# 87128-25G), p-tyramine (Sigma Aldrich, catalog# T2879-1G), DL-3,4-Dihydroxymandelic acid (Carbo Synth, catalog# FD22118), protocatechuic Acid (Millipore Sigma, catalog# 37580-25G-F), DL norepinephrine (Millipore Sigma, catalog# A7256-1G), L-norepinephrine (Matrix Scientific, catalog# 037592-500MG) L-epinephrine (Alfa Aesar, catalog# L04911.06), DL-epinephrine (Sigma Aldrich, catalog# E4642-5G), 3,4-dihydroxyphenylacetic acid (Millipore Sigma, catalog# 850217-1G), 3,4-dihydroxyhydrocinnamic acid (hydrocaffeic acid) (Millipore Sigma, catalog# 102601-10G), caffeic acid (Millipore Sigma, catalog# C0625-2G), (+)-catechin hydrate (Millimore Sigma, catalog# C1251-5G), (+/-)-catechin hydrate (Millipore Sigma, catalog# C1788-500MG), (-)-Epicatechin (Millipore Sigma, catalog# E1753-1G), L-(-)-a-Methyldopa (Chemcruz, catalog# sc-203092), 2,3-dihydroxybenzoic acid (Millipore Sigma, catalog# 126209-5G), R-(-)-apomorphine hydrochloride hemihydrate (Sigma Aldrich, catalog# A4393-100MG), hydroxytyrosol (Ava Chem Scientific, catalog# 2528), enterobactin (generous gift from Prof. Elizabeth Nolan, MIT), fenoldopam mesylate (Sigma Aldrich, catalog# SML0198-10MG), 5-hydroxydopamine (Sigma Aldrich, catalog# 151564-100G), 6-hydroxydopamine (Sigma Aldrich, catalog # H4381-100MG), 3-methoxytyramine (Sigma Aldrich, catalog# M4251-100MG), 3,4-dihydroxybenzylamine (Sigma Aldrich, catalog# 858781-250MG), N-methyl dopamine (Santa Cruz Biotechnology, catalog# sc-358430A), 4-(2-aminoethyl)benzene-1,3-diol (Enamine, catalog # EN300-65185), m-tyramine (Chemcruz, catalog# sc-255257), 3-hydroxyphenylacetic acid (Sigma Aldrich, catalog# H49901-5G), 3-hydroxyphenylpropionic acid (Toronto Research Chemicals, catalog# H940090), L-dopa (Oakwood Chemical, catalog# 358380-
25g), dopamine (Sigma-Aldrich, catalog# PHR1090-1G, or Millipore Sigma, catalog# H8502-25G), m-tyramine (Santa Cruz Biotechnology, catalog# sc-255257), carbidopa (Sigma-Aldrich, catalog# PHR1655-1G), L-arginine (Sigma-Aldrich, catalog# A5006-100G), sodium molybdate (Sigma-Aldrich, catalog# 243655-100G), sodium tungstate (72069-25G), SIGMAFAST protease inhibitor tablets (Sigma-Aldrich, catalog#: S8830), benzyl viologen (Sigma-Aldrich, catalog# 271845-250mg), methyl viologen (Sigma-Aldrich, catalog# 856177-1g), diquat (Sigma-Aldrich, catalog# 45422-250mg), sodium dithionite (Sigma-Aldrich, catalog# 157953-5G), 3,4-dihydroxyphenylacetic acid (ring-d3, 2,2-d2, 98%) (Cambridge Isotope Laboratories, catalog #DLM-2499-0.01), dopamine HCl (1,1,2,2-d4, 97-98%) (Cambridge Isotope Laboratories, catalog #DLM-2498-0.1), caffeic acid (Sigma-Aldrich, catalog# C0625-2G). LC-MS grade acetonitrile and methanol for LC-MS analyses were purchased from Honeywell Burdick & Jackson or Sigma-Aldrich. Brain Heart Infusion (BHI) broth was purchased from Beckton Dickinson (catalog# 211060) or from VWR (catalog# 95021-488).

All bacterial culturing work was performed in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen, 10% carbon dioxide, and nitrogen as the balance, unless otherwise noted. Hungate tubes were used for anaerobic culturing unless otherwise noted (Chemglass, catalog# CLS-4209-01). All lysate work and biochemical experiments were performed in an anaerobic chamber (Coy Laboratory Products) situated in a cold room at 4 °C under an atmosphere of 10% hydrogen and nitrogen as the balance. Gut Actinobacterial strains were grown on BHI containing 1% arginine (w/v) to obtain isolated colonies for culturing.

All genomic DNA (gDNA) was extracted from bacterial cultures using the DNeasy UltraClean Microbial Kit (Qiagen, catalog # 12224-50) according to the manufacturer’s protocol.
4.4.2 LC-MS methods

Method A: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801). The flow rate was 0.5 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 0% B isocratic, 2-9 min: 0-10% B, 9-11 min: 10-95% B, 11-15 min: 95% B isocratic, 15-18 min: 95-0% B, 18-21 min: 0% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of dopamine (precursor ion \(m/z = 154.3\), daughter ion \(m/z = 137.3\)), and tyramine (precursor ion \(m/z = 138.3\), daughter ion \(m/z = 121.3\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method B: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of trihydroxydopamine (precursor ion \(m/z = 170.3\), daughter ion \(m/z = 153.3\)), dopamine (precursor ion \(m/z = 154.3\), daughter ion \(m/z = 137.3\)), phenylethylamine (precursor ion \(m/z = 122.3\), daughter ion \(m/z = 105.2\)), and tyramine (precursor ion \(m/z = 138.3\), daughter ion \(m/z = 121.3\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method C: Samples were analyzed using an Agilent technologies 6530 Accurate-Mass Q-TOF LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog
Method D: Samples were analyzed using an Agilent technologies 6530 Accurate-Mass Q-TOF LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801). The flow rate was 0.4 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 5% B isocratic, 2-25 min: 0-95% B, 25-30 min: 95% B isocratic, 30-40 min: 95-5% B. For the MS detection, the ESI mass spectra data were recorded in positive mode for a mass range of m/z 50 to 3000. A mass window of ± 0.005 Da was used to extract the ion of [M+H].

Method E: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of catechin (precursor ion m/z = 289.2, daughter ion m/z = 109.1), benzyl ether reduced catechin (precursor ion m/z = 291.2, daughter ion m/z = 123.1), benzyl ether reduced, dehydroxylated catechin (precursor ion m/z = 275.2, daughter ion m/z = 107.1) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative
Method F: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of hydrocaffeic acid (precursor ion \(m/z = 181.2\), daughter ion \(m/z = 137.2\)), hydroxyphenylpropionic acid (precursor ion \(m/z = 165.1\), daughter ion \(m/z = 121.2\)), DOPAC (precursor ion \(m/z = 167.2\), daughter ion \(m/z = 123.2\)), and hydroxyphenylacetic acid (precursor ion \(m/z = 151.3\), daughter ion \(m/z = 107.3\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

Method G: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim polar advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 275 °C, and the masses of norepinephrine (precursor ion \(m/z = 170.1\), daughter ion \(m/z = 152.1\)) and octopamine (precursor ion \(m/z = 154.2\), daughter ion \(m/z = 136.1\)) were monitored at a collision energy of 5 mV and fragmentor setting of 135 in positive MRM mode.

Method H: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile
phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 275 °C, and the masses of caffeic acid (precursor ion m/z = 179.2, daughter ion m/z = 135.2) and coumaric acid (precursor ion m/z = 163.3, daughter ion m/z =119.2) were monitored at a collision energy of 5 mV and fragmentor setting of 135 in negative MRM mode.

Method I: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of dihydroxybenzoic acid (precursor ion m/z = 153.1, daughter ion m/z = 137.1) and hydroxybenzoic acid (precursor ion m/z = 137.1, daughter ion m/z = 93.2) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

Method J: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of norepinephrine (precursor ion m/z = 184.1, daughter ion m/z = 166.1) and dehydroxynorepinephrine (precursor ion m/z = 168.1, daughter ion m/z =
150.1) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method K: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of dihydroxybenzylamine (precursor ion \(m/z = 140.3\), daughter ion \(m/z = 123.2\)) and hydroxybenzylamine (precursor ion \(m/z = 124.3\), daughter ion \(m/z = 107.2\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method L: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of 3-aminotyramine (precursor ion \(m/z = 153.3\), daughter ion \(m/z = 136.2\)) and 3-aminophenylethylamine (precursor ion \(m/z = 137.3\), daughter ion \(m/z = 120.2\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method M: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile
phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses 3-methoxytyramine (precursor ion $m/z = 151.1$, daughter ion $m/z = 91.1$) and 3-methoxyphenylethylamine (precursor ion $m/z = 135.1$, daughter ion $m/z = 75.1$) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method N: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min$^{-1}$ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses 3-hydroxytyrosol (precursor ion $m/z = 153.2$, daughter ion $m/z = 123.1$) and tyrosol (precursor ion $m/z = 137.2$, daughter ion $m/z = 107.1$) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

Method O: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.15 mL min$^{-1}$ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of d5-DOPAC (precursor ion $m/z = 172.2$, daughter ion $m/z = 128.2$) and d5-hydroxyphenylacetic acid (precursor ion $m/z = 156.3$, daughter ion $m/z = 113.3$) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.
Method P: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 µM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of d4-dopamine (precursor ion m/z = 158.3, daughter ion m/z = 141.3), and d4-tyramine (precursor ion m/z = 142.3, daughter ion m/z = 125.3) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method Q: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 µM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of caffeic acid (precursor ion m/z = 179.2, daughter ion m/z = 135.2), coumaric acid (precursor ion m/z = 163.1, daughter ion m/z = 119.2), DOPAC were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

4.4.3 Colorimetric assay for catechol detection

Our colorimetric assay for dopamine dehydroxylation was based on the Arnow test (103). Briefly, 50 µL of 0.5 M HCl was added to 50 µL of culture supernatant. After mixing, 50 µL of an aqueous solution containing both sodium molybdate and sodium nitrite (0.1 g/mL each) was added, which produced a yellow color. Finally, 50 µL of 1 M NaOH was added followed by pipetting up and down to mix. This allowed the characteristic pink color to develop. Absorbance
was measured at 500 nm immediately using a Synergy HTX Multi-Mode Microplate Reader (BioTek) or SPECTROstar Nano (BMG LABTECH).

4.4.4 Anaerobic activity-based purification of E. lenta A2 dopamine dehydroxylase

Protein purification: Experiments were performed as described previously (see Chapter 3), with minor modifications. All procedures were carried out under strictly anaerobic conditions at 4 °C. Procedures outside the anaerobic chamber were performed in tightly sealed containers to prevent oxygen contamination. First, E. lenta A2 starter cultures were inoculated from single colonies into liquid BHI medium and were grown for 30 hours. Starter cultures were diluted 1:100 into 5 L of BHI medium containing 1% arginine and 10 mM formate and grown anaerobically at 37 °C for 16 hours. Dopamine was added as a solid to a final concentration of 0.5 mM in the cultures. Cells were pelleted in 5 separate 1 L bottles by centrifugation (6000 rpm, 15 mins), and each pellet was resuspended in 20 mL of 20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail. Resuspended cells were then lysed using two rounds of sonication in an anaerobic chamber (Branson Sonifier 450, 2 min total, 10 sec on, 40 sec off, 25% amplitude). The lysates were then clarified by centrifugation (10800 rpm, 15 mins), and the soluble fractions were subjected to two rounds of ammonium sulfate precipitation. During the precipitation, three different tubes each containing 40 mL total clarified lysate were precipitated in parallel. Solid ammonium sulfate was first dissolved in these clarified lysates to a final concentration of 30% (w/v), and the lysates were left for 1 hour and 20 minutes followed by centrifugation to pellet the precipitates (4000 rpm, 15 mins). The supernatant was saved, and the pellet was discarded. The supernatant was mixed with additional solid ammonium sulfate to achieve a final concentration of 40% (w/v) and left for 1 hour and 20 minutes. Following centrifugation (4000 rpm, 15 mins) and removal of supernatant, each pellet containing the precipitated proteins was re-dissolved in 20 mL.
20 mM Tris pH 8 containing 0.5 M ammonium sulfate. The re-dissolved pellets were combined and centrifuged to remove particulates (10800 rpm, 15 mins). The resulting 60 mL solution was injected onto an FPLC (Bio-Rad BioLogic DuoFlow System equipped with GE Life Sciences DynaLoop90) for hydrophobic interaction chromatography (HIC) using 5 x 1mL HiTrap Phenyl HP columns (GE Life Sciences, catalog# 17135101). Fractions were eluted with a gradient of 0.5 M to 0 M ammonium sulfate (in 20 mM Tris pH 8) at a flow rate of 1 mL/min and were tested for activity using the assay described below. The majority of the dopamine dehydroxylase activity eluted around 0.05 M-0.1 M ammonium sulfate. Active fractions displaying >50% conversion of dopamine were combined and injected onto the FPLC system described above for anion exchange chromatography using a UNO Q1 column (Bio-Rad, catalog# 720-0001) at a flow-rate of 1 mL/min. Fractions were eluted using a gradient of 0 to 1 M NaCl in 20 mM Tris pH 8 and were tested for activity. The majority of the dopamine dehydroxylase activity eluted around 250 mM NaCl. Active fractions were combined and concentrated 20-fold using a spin concentrator with a 5 kDa cutoff (4000 rpm centrifugation speed). 250 µL of the concentrate was injected onto FPLC for size exclusion chromatography using an Enrich 24 mL column (Enrich SEC 650, 10*300 column, Bio-Rad, catalog# 780-1650). Fractions were eluted over a 26 mL volume run isocratically in 20 mM Tris pH 8 containing 250 mM NaCl, collected, and were subjected to activity assays. Active fractions were then combined and analyzing SDS-PAGE to assess the presence of protein. Absorbance at 280 nm was used to determine the protein concentration, using a predicted extinction coefficient of 317735 M⁻¹ cm⁻¹ for the dopamine dehydroxylase (determined by the PROTPARAM tool: https://web.expasy.org/protparam/).

Activity assays during protein purification: 50 µL aliquots of the combined active fractions from FPLC runs were mixed, in the following order, with 1 µL electron donors (final concentration
1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water), 2 µL sodium dithionite (2 mM final concentration, dissolved in water), and 1 µL substrate (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for 12–14 hours to allow dopamine dehydroxylation to proceed, followed by assessment of activity using the colorimetric assay for catechol detection. Due to the inability of Dadh to tolerate freeze-thawing even in the presence of glycerol, the natively purified enzyme was always immediately used for enzyme assays.

### 4.4.5 Assays of the *E. lenta* A2 dopamine dehydroxylase substrate scope

Active fractions from the size exclusion chromatography step described above were diluted in 20 mM Tris pH 8 containing 250 mM NaCl to a final enzyme concentration of 0.1 µM. The enzyme mixture was transferred into the wells of a 96 well plate, for a final volume of 50 µL in each well (VWR, catalog# 82006-636). 1 µL of substrate (in water, or 50:50 water:DMF for caffeic acid and catechin substrates) was then added at a final concentration of 500 µM. Following this, 1 µL of a solution containing electron donors (final concentration 1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water) and 2 µL of sodium dithionite (2 mM final concentration, dissolved in water) were added. The resulting solution was mixed by pipetting and the 96-well plate was then sealed tightly with an aluminum seal. The enzyme assay mixtures were left at room temperature in an anaerobic chamber for 22 hours to allow dehydroxylation to proceed. The enzyme reaction mixtures were quenched by bringing the samples out of the anaerobic chamber and freezing at –20 °C. These mixtures were then diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS. For the screen with physiologically relevant catechol substrates, samples containing caffeic acid were analyzed using Method H, samples containing hydrocaffeic acid and DOPAC were analyzed using Method F, catechin was analyzed
using Method E, samples containing protocatechuic acid were analyzed using Method I, samples containing epinephrine were analyzed using Method J, samples containing norepinephrine were analyzed using Method G, and samples containing ellagic acid were analyzed using Method D. For the screen with dopamine analogs, all samples containing monohydroxylated, dihydroxylated, and trihydroxylated phenylethylamine analogs were analyzed using method B, samples containing N-methyldopamine were analyzed using Method C, samples containing methoxytyramine were analyzed using Method M, samples containing dihydroxybenzylamine were analyzed using Method K, samples containing hydroxytyrosol were analyzed using Method N, and samples containing aminotyramine were analyzed using Method L. Reactions were also analyzed by the colorimetric assay for catechol detection and high-resolution LC-MS (QTOF) using methods C and D to verify metabolism.

4.4.6 Metabolism of dopamine analogs by E. lenta A2 cells

Cells were cultured in 96-well plates and all experiments were performed anaerobically. The strains screened for dopamine dehydroxylation have been described (see Chapter 3). E. lenta A2 was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:10 in triplicate into 200 µL of fresh BHI medium containing 500 µM substrate (p-tyramine, dopamine, 3,4-dihydroxybenzylamine, or DL-norepinephrine). These cultures were grown for 48 hours at 37 °C. Cultures were harvested by centrifugation at 4000 rpm for 10 minutes, and the supernatants were diluted 1:10 with LC-MS grade methanol. Samples containing dopamine or p-tyramine were analyzed using Method B, samples containing norepinephrine were analyzed using Method G, samples containing dihydroxybenzylamine were analyzed using Method K. Cultures were also analyzed by the colorimetric assay to verify the results.
4.4.7 RNA-sequencing experiments with *E. lenta*

We repeated the setup previously used in the RNA-sequencing experiment with dopamine (see Chapter 3). Turbid 48-hour starter cultures of *E. lenta* in BHI medium were inoculated 1:100 into 5 mL of BHI medium containing 1% arginine and 10 mM formate, and cultures were grown at 37 °C anaerobically. When the cultures reached OD$_{600}$=0.200, hydrocaffeic acid, (+)-catechin, $p$-tyramine, 3,4-dihydroxybenzylamine, DL-norepinephrine, or caffeic acid were added at final concentrations of 500 µM to triplicate cultures. All compounds except for (+)-catechin and caffeic acid were dissolved in water; (+)-catechin and caffeic acid were dissolved in DMF. Control cultures contained vehicle (water or DMF). Cultures were harvested when they reached OD$_{600}$=0.500. They were centrifuged for 15 minutes at 4000 rpm, and cell pellets were re-suspended in 500 µL Trizol reagent (ThermoFisher, catalog#: 15596026).

All work following freezing of bacterial pellets was performed by the Microbial Omics Core (MOC) at the Broad Institute. Total RNA was isolated by first bead beating to lyse cells and then using the Zymo Research Direct-Zol RNA MiniPrep Plus kit (Catalog # R2070) according to the manufacturer’s protocol. Illumina cDNA libraries were generated using a modified version of the RNAtag-Seq protocol (104). Briefly, 500 ng of total RNA was fragmented, depleted of genomic DNA, and dephosphorylated prior to its ligation to DNA adapters carrying 5'-AN8-3' barcodes with a 5' phosphate and a 3' blocking group. Barcoded RNAs were pooled and depleted of rRNA using the RiboZero rRNA depletion kit (Epicentre). These pools of barcoded RNAs were converted to Illumina cDNA libraries in 3 main steps: (i) reverse transcription of the RNA using a primer designed to the constant region of the barcoded adaptor; (ii) addition of a second adapter on the 3’ end of the cDNA during reverse transcription using SmartScribe RT (Clonetech) as described (104); (iii) PCR amplification using primers that target the constant regions of the 3’ and
5’ ligated adaptors and contain the full sequence of the Illumina sequencing adaptors. cDNA libraries were sequenced on Illumina HiSeq 2500. For the analysis of RNAtag-Seq data, reads from each sample in the pool were identified based on their associated barcode using custom scripts, and up to 1 mismatch in the barcode was allowed with the caveat that it did not enable assignment to more than one barcode. Barcode sequences were removed from the first read as were terminal G’s from the second read that may have been added by SMARTScribe during template switching. Reads were aligned to the *Eggerthella lenta* genomes (strains A2 or 28b) using BWA (105) and read counts were assigned to genes and other genomic features using custom scripts. Differential expression analysis was conducted with DESeq2 (106) and/or edgeR (107).

4.4.8 Time course of DOPAC metabolism by *G. pamelaeae* 3C

*Method 1* (compound added at mid-exponential phase): Turbid 48-hour starter cultures of *G. pamelaeae* 3C grown in of BHI medium were inoculated 1:100 into triplicate Hungate tubes containing 10 mL of BHI medium with 10 mM formate. When cultures reached OD$_{600}$=0.130, DOPAC (0.5 mM final) or vehicle (water) was added to the cultures. The cultures were then grown at 37 °C anaerobically and aliquots were withdrawn periodically, harvested by centrifugation, and the supernatant was analyzed for metabolism using the colorimetric assay. *Method 2* (compound added at the beginning of growth): Turbid 48-hour starter cultures of *G. pamelaeae* 3C grown in BHI medium were inoculated 1:100 into triplicate Hungate tubes containing 10 mL BHI medium with 10 mM formate and DOPAC (0.5 mM final) or vehicle (water). These cultures were then left to grow at 37 °C anaerobically and aliquots were withdrawn periodically, harvested by centrifugation, and the supernatant was analyzed for metabolism using the colorimetric assay.
4.4.9 RNA-sequencing experiments with *G. pamelaeae 3C*

*Method 1* (compound added during exponential phase): Turbid 48-hour starter cultures of *G. pamelaeae 3C* grown in BHI medium were inoculated 1:100 into triplicate Hungate tubes containing 20 mL of BHI medium with 10 mM formate. When cultures reached $\text{OD}_{600}=0.110$, DOPAC (0.5 mM final) or vehicle (water) was added to the cultures. The cultures were then grown at 37 °C anaerobically and harvested when they reached $\text{OD}_{600}=0.185$. They were centrifuged, and cell pellets were re-suspended in 500 µL Trizol reagent (ThermoFisher, catalog#: 15596026).

*Method 2* (compound added at the beginning of growth): Turbid 48-hour starter cultures of *G. pamelaeae 3C* grown in BHI medium were inoculated 1:100 into triplicate Hungate tubes containing 20 mL of BHI medium with 10 mM formate and DOPAC (0.5 mM final) or vehicle (water). These cultures were then left to grow at 37 °C anaerobically. When cultures reached $\text{OD}_{600}=0.110$, they were harvested. They were centrifuged, and cell pellets were re-suspended in 500 µL Trizol reagent (ThermoFisher, catalog#: 15596026). RNA extraction and sequencing: this was performed using exactly the same setup as described above, except the reads were aligned to the genome of *Gordonibacter pamelaeae 3C*.

4.4.10 Growth of *E. lenta A2* in BHI medium with and without dopamine

Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *E. lenta A2* was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 5 mL of BHI medium containing either 0.5 mM dopamine or vehicle. Growth was assessed by measuring the optical density at 600 nm using a Genesys 20 spectrophotometer (Thermo Scientific).
4.4.11 Preparation of basal medium lacking electron acceptors

The medium was prepared as described previously, with minor modifications (see chapter 3 of this thesis). A 100-fold stock solution of salts was first prepared by dissolving 100 g NaCl, 50 g MgCl₂•6H₂O, 20 g KH₂PO₄, 30 g NH₄Cl, 30 g KCl, 1.5 g CaCl₂ x 2H₂O in 1 L of water. Then, 10 mL of this solution was added to 1 L of water containing 1 g yeast extract (Beckton Dickinson #288260), 1 g tryptone (Beckton Dickinson #21175), and 0.25 mL of 0.1% resazurin (dissolved in MilliQ water). This medium was autoclaved. Following autoclaving, the medium was left to cool for 15 minutes in an atmosphere of air (outside the anaerobic chamber). After cooling, the following components were added using sterile technique: 10 mL of ATCC Trace element mix (ATCC, catalog# MD-TMS), 10 mL of Vitamin Supplement (ATCC, catalog# MD-VS), solid NaHCO₃ (SIGMA, 2.52 g, to give 30 mM) and solid L-cysteine HCl (SIGMA, 63 mg, to give 0.4 mM). The medium had a final pH of 7.2-7.3. The medium was then sparged with nitrogen gas for 30 minutes and was brought into the anaerobic chamber to equilibrate for at least 30 hours prior to use. In all experiments utilizing the basal medium, except for those experiments performed with Gordonibacter pamelaeae 3C or the screen for catechol metabolism by mammalian gut microbiota samples, sodium acetate was added at a final concentration of 10 mM at the time of inoculation. In experiments performed with Gordonibacter pamelaeae 3C, sodium formate was added at a final concentration of 10 mM. In the ex vivo experiments with the mammalian gut microbiota, neither acetate nor formate were added to the basal medium.

4.4.12 Growth of single E. lenta strains in basal medium

Strains were cultured in Hungate tubes and all experiments were performed anaerobically. E. lenta strains were inoculated from single colonies into 10 mL of BHI liquid medium and grown for 48-72 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate
into 5 mL of basal medium containing 10 mM acetate and either 1 mM dopamine (in water) or vehicle (water). If applicable, molybdate (0.5 mM), tungstate (0.5 mM), DMSO (14 mM), (+)-catechin (1 mM), hydrocaffeic acid (1 mM), or nitrate (1 mM) were added at the time of inoculation. Cultures were grown anaerobically for 36-72 hours at 37 °C. Endpoint growth was assessed by measuring the optical density at 600 nm using a Genesys 20 spectrophotometer (Thermo Scientific). Catechol dehydroxylation was assessed at the end of growth in culture supernatants using the colorimetric method.

**4.4.13 Growth of *G. pamelaeae* 3C strains in basal medium**

Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *G. pamelaeae* 3C was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 5 mL of basal medium containing 10 mM acetate or 10 mM formate and either 1 mM DOPAC (in water) or vehicle (water). Cultures were grown anaerobically for 72 hours at 37 °C. Endpoint growth was assessed by measuring the optical density at 600 nm using a Genesys 20 spectrophotometer (Thermo Scientific). Growth was also measured by a spot plating method. Catechol dehydroxylation was assessed at the end of growth in culture supernatants using the colorimetric method.

**4.4.14 Competition of *E. lenta* strains in basal medium**

Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *E. lenta* strains A2 (Tet resistant metabolizer, white colony color), DSM2243 (Tet sensitive non-metabolizer, pink colony color), W1BHI6 (Tet resistant non-metabolizer, white colony color), and Valencia (Tet sensitive metabolizer, white colony color) were inoculated from single colonies into 10 mL BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. For
the competition experiment, 50 µL of each starter culture of the two competing strains was combined in triplicate into 5 mL of basal medium containing 10 mM acetate and either 1 mM dopamine or vehicle (water). The competing pairs were either A2/DSM2243 or W1BHI6/Valencia. Following inoculation, cultures were grown anaerobically for 72 hours at 37 °C. At the end of the incubation, growth of *E. lenta* was assessed. Cultures were serially diluted in PBS under anaerobic conditions, and 8 uL of each serial dilution (10⁻¹ through 10⁻⁷) was plated onto BHI agar plates containing 1% arginine (w/v) with and without 10 µg/mL tetracycline using a spot plating method. Plates were grown at 37 °C for 72 hours followed by counting of colonies. To calculate the proportion of metabolizer in the A2/DSM2243 competition experiment, we selected a serial dilution where distinct colonies were clearly visible (10⁻⁴). We then counted the number of white (A2) and pink colonies (DSM2243) growing on BHI 1% arginine agar plates lacking tetracycline. To calculate the proportion of metabolizer in the W1BHI6/Valencia competition experiment, we performed a serial dilution where distinct colonies were clearly visible (10⁻⁴-10⁻⁵) and counted the number of colonies growing on the BHI 1% arginine tetracycline agar plates (W1BHI6) as well as the colonies growing on the BHI 1% arginine agar plates (Both Valencia and W1BHI6). To calculate the number of metabolizer (Valencia) colonies, we subtracted the number of tetracycline resistant colonies from the colonies on the plate lacking tetracycline.

**4.4.15 Human fecal samples used in this study**

The human fecal samples used in this study have been previously described (see Chapter 3 of this thesis). To prepare them for culturing, all samples were resuspended anaerobically in anaerobic PBS at a final concentration of 0.1 g/mL. The mixture was vortexed to produce a homogenous slurry and was then left for 30 minutes to let particulates settle. Aliquots of the
supernatant were dissolved 50:50 with 40% glycerol and flash-frozen in liquid nitrogen, creating slurries that were used for anaerobic culturing of human fecal samples. Slurries were stored at −80 °C and were thawed anaerobically at room temperature at the time of use.

4.4.16 Growth of human fecal samples in basal medium with dopamine

Fecal slurries from n=24 unrelated humans were diluted 1:100 into two different Hungate tubes containing 5 mL of basal medium with 10 mM acetate and either 1 mM dopamine or vehicle (water). These fecal microbiota cultures were grown anaerobically for 72 hours at 37 °C. Metabolism was then assessed in culture supernatants using the colorimetric method. In addition, cultures were spun down and the total community gDNA was extracted from the entire 5 mL of culture for downstream PCR and qPCR assays as detailed below.

4.4.17 qPCR assays for E. lenta and dadh abundance in human fecal samples grown in basal medium with and without dopamine

gDNA was extracted from the culture pellets generated in the experiments described above (“Growth of fecal samples in basal medium with dopamine”) using the DNeasy UltraClean Microbial Kit. The extracted DNA from each culture was used for qPCR assays containing 10 µL of iTaq Universal SYBRgreen Supermix (Bio-rad, catalog 3: 1725121), 7 µL of water, and 10 µM each of forward and reverse primers. PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 95 °C for 5 minutes 34 cycles of 95 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min. The program ended with a final extension at 34 °C for 5 mins. The primers used were: 16S rRNA primers for E. lenta (108): 5’ CAGCAGGGAAGAAATTCGAC 3’ and 5’ TTGAGCCCTCGGATTAGAG 3’; primers for dopamine dehydroxylase: 5’ GAGATCTGCGTCCACCGTCA 3’ and 5’ AGTGGAAGTACACCGGGAT 3’.
4.4.18 Amplification of full-length *dadh* and sequencing of the SNP at position 506 from human fecal samples grown in basal medium with and without dopamine

gDNA was extracted from the culture pellets generated in the experiments described above (“Growth of fecal samples in basal medium with dopamine”) using the DNeasy UltraClean Microbial Kit. The extracted DNA from each culture was used for PCR assays containing 10 µL of Phusion High-Fidelity PCR Master mix with HF buffer (NEB, catalog# M0531L), 7 µL of water, and 10 µM each of forward and reverse primers. The primers used to amplify the full-length dopamine dehydroxylase from these samples were 5’ ATGGGTAACCTGACCATG 3’ and 5’ TTACTCCCTCCCTCGTA 3’. PCR was performed on a C1000 Touch Thermocycler (Bio-Rad), using the following program: initial denaturation at 98 ºC for 30 s, 34 cycles of 98 ºC for 10 s, 61 ºC for 15 s, 72 ºC for 2.5 mins. The program ended with a final extension at 72 ºC for 5 mins. Amplicons were purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, catalog# 28-9034-70) and were sequenced using Sanger sequencing (Eton Biosciences) for the region containing the SNP at position 506 using primers 5’ GGGGTGTCCATGTTGCCGGT 3’ and 5’ ACCGGCTACGGCAACGGC 3’. Sequence chromatograms were analyzed in Ape Plasmid Editor (version 2.0.47), and the single nucleotide polymorphism (SNP) at position 506 was called by visual inspection compared to results obtained from control cultures of *E. lenta* strains.

4.4.19 Screen of gut Actinobacteria for metabolism of catechols

This procedure was performed in an anaerobic chamber (Coy Laboratory Products, atmospheric conditions: 20% CO₂, 2-2.5% H₂, and the balance N₂)—equilibrating media and consumables to the atmosphere prior to use—until centrifugation, which was performed using a benchtop centrifuge. The 96-well plates used in this experiment were purchased from VWR.
Into the wells of flat-bottom 96-well plates, 100 μL of BHI medium supplemented with L-cysteine-HCl (0.05%, w/v), L-arginine (1%, w/v), and sodium formate (10 mM) (referred to here as BHI++) were aliquoted. Seed cultures were prepared by inoculating wells, in triplicate, with Actinobacterial strains that were cultured on BHI++ agar plates. Additional wells served as sterile controls. Plates were sealed with tape and incubated at 37 °C for 12 to 18 hours to afford dense cultures. Next, 99 μL of BHI++ medium containing 500 μM of compound were aliquoted into the wells of a 96-well plate. To these wells, 1 μL of dense seed culture (or sterile control) was added. Plates were sealed and incubated at 37 °C for 24 or 48 hours. Plates were then centrifuged at 2000 rpm for 10 min at 4 °C, and the supernatant was aspirated and transferred to a fresh 96-well plate. An aliquot (35 μL) of supernatant was then immediately screened using the catechol colorimetric assay (described above in “Colorimetric assay for catechol detection”). Absorbance was immediately measured at 500 nm using a plate reader (Spectrostar Nano, BMG LABTECH). A standard curve (2-fold serial dilutions, 1000-15.6 μM in BHI++) was simultaneously prepared, developed, and analyzed using the conditions listed above. The catechol concentrations in bacterial cultures were normalized to the sterile control. All cultures except for ellagic acid were prepared in this way and analyzed by the colorimetric assay. However, in incubations with ellagic acid, strains were grown in BHI medium without any additives, and culture supernatants were analyzed by LC-MS using method D. For LC-MS/MS spot-checking of the colorimetric screen, cultures from strains incubated with caffeic acid were analyzed using Method Q and Method F, cultures from strains incubated with DOPAC were analyzed using Method F, and cultures from strains incubated with norepinephrine were analyzed using Method G. All these initial screens, including culturing and the colorimetric method, were performed by Paola Nol (Bess lab, UC Irvine). Vayu Maini Rekdal performed all LC-MS analysis.
To confirm metabolism of (+)-catechin, DOPAC, and hydrocaffeic acid, the incubations were repeated following the same procedure with minor modifications. Strains were grown in BHI medium for 48 hours anaerobically at 37 °C. Cultures were harvested by centrifugation and were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to a 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method E for samples containing catechin and Method F for samples containing hydrocaffeic acid and DOPAC. Following this screen, select strains were re-cultured to confirm absence/presence of metabolism.

The following stock solutions were used in the screens: dihydroxymandelic acid (50 mM in water), dopamine (50 mM in water), protocatechuic acid (50 mM in ethanol), L-dopa (50 mM in 0.5 M HCl), norepinephrine (50 mM in 0.5 M HCl), epinephrine (50 mM in 0.5 M HCl), DOPAC (50 mM in 0.5 M HCl), Hydrocaffeic acid (50 mM in ethanol), caffeic acid (50 mM in ethanol), (+)-catechin (50 mM in ethanol), (+/-)-catechin (50 mM in ethanol), (-)-epicatechin (50 mM in DMSO), methyldopa (solid dissolved directly into the media at 0.5 mM final concentration), carbidopa (solid dissolved directly into the media at 0.5 mM final concentration), dihydroxybenzoic acid (50 mM in methanol), hydroxytyrosol (50 mM in water), enterobactin (10 mM in DMSO), and apomorphine (50 mM in DMSO).

4.4.20 Confirmation of dehydroxylation of (+)-catechin and hydrocaffeic acid by E. lenta A2

Cells were cultured in Hungate tubes and all experiments were performed anaerobically. E. lenta A2 was inoculated from a single colony into 10 mL BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 5 mL of BHI medium containing either 0.5 mM hydrocaffeic acid (in water), 0.5 mM (+)-catechin (in
DMF), or vehicle (water or DMF). After 48 hours of anaerobic growth at 37 °C, cultures were harvested by centrifugation and were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to a 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method E for samples containing catechin and Method F for samples containing hydrocaffeic acid.

**4.4.21 Confirmation of dehydroxylation of DOPAC by *Gordonibacter pamelaeeae 3C***

Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *G. pamelaeeae 3C* was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 5 mL of BHI medium containing 10 mM formate and 0.5 mM DOPAC or vehicle (water). After 72 hours of anaerobic growth at 37 °C, the cultures were harvested by centrifugation and were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument and analyzed using Method F for DOPAC.

**4.4.22 PBS resuspension assays for inducibility and oxygen sensitivity of catechol dehydroxylases**

Assays with *E. lenta A2*: Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *E. lenta A2* was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were
diluted 1:100 in triplicate into 10 mL of BHI medium containing 1% arginine and 10 mM formate and either 0.5 mM dopamine (in water), 0.5 mM hydrocaffeic acid (in water), 0.5 mM (+)-catechin (in DMF), or vehicle (water or DMF). After 18 hours of anaerobic growth at 37 °C, cultures had reached an OD$_{600}$ of 0.700 and were harvested by centrifugation (4000 rpm, 15 minutes). The bacterial pellets were resuspended anaerobically in 10 mL of pre-reduced PBS to wash the cells, followed by an additional round of centrifugation to pellet the washed cells (4000 rpm, 15 minutes). The cells were then resuspended in 5 mL of pre-reduced PBS. 0.1 mL aliquots of this resuspension were transferred to Eppendorf tubes containing either vehicle or 0.5 mM catechol substrate. The samples were vortexed briefly and incubated anaerobically at room temperature for 20 hours to allow for metabolism to proceed. To assess the impact of oxygen on the metabolism of catechols, 0.1 mL of the PBS resuspension in an Eppendorf tube was brought outside the anaerobic chamber, followed by addition of 0.5 mM substrate in the presence of atmospheric oxygen. The samples were vortexed briefly and were incubated at room temperature for 20 hours. Samples were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method B for samples containing dopamine, Method E for samples containing catechin and Method F for samples containing hydrocaffeic acid.

Assays with *Gordonibacter pamelaeae* 3C: Cells were cultured in Hungate tubes and all experiments were performed anaerobically. *G. pamelaeae* 3C was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 10 mL of BHI medium containing 10 mM formate and
either 0.5 mM DOPAC or vehicle. After 18 hours of anaerobic growth at 37 °C, cultures had reached an OD$_{600}$ of 0.180 and were harvested by centrifugation (4000 rpm, 15 minutes). The bacterial pellets were resuspended anaerobically in 10 mL of pre-reduced PBS to wash the cells, followed by an additional round of centrifugation to pellet the washed cells. The cells were then resuspended in 5 mL of pre-reduced PBS. 0.1 mL aliquots of this resuspension were transferred to Eppendorf tubes containing either vehicle or 0.5 mM catechol substrate. The samples were vortexed briefly and incubated anaerobically at room temperature for 20 hours to allow for metabolism to proceed. To assess the impact of oxygen on DOPAC metabolism, 0.1 mL of the PBS resuspension in an Eppendorf tube was brought outside the anaerobic chamber, followed by addition of 0.5 mM substrate in the presence of atmospheric oxygen. The samples were vortexed briefly and were incubated at room temperature for 20 hours. Samples were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method F.

4.4.23 Effect of tungstate on growth and catechol dehydroxylation by *E. lenta* A2 and *G. pamelaeae* 3C

Assays with *E. lenta* A2: Starter cultures of *E. lenta* A2 were grown over 48 hours in 10 mL of BHI medium and then inoculated 1:100 into 200 µL of BHI medium containing either 500 µM dopamine (in water), 500 µM (+)-catechin (in DMF), or 500 µM hydrocaffeic acid (in water), and either sodium tungstate (0.5 mM, in water), sodium molybdate (0.5 mM, in water), or vehicle (water or DMF). Cultures were grown for 48 hours anaerobically at 37 °C and were harvested by centrifugation. Supernatants were dissolved 1:10 in LC-MS grade methanol and analyzed using
LC-MS/MS Methods B, E, or F as described above. Experiments were performed anaerobically, and cultures were grown in 96-well plates (VWR, catalog# 29442-054).

**Assays with *Gordonibacter pamelaeae* 3C:** Starter cultures of *G. pamelaeae* 3C were grown over 48 hours in 10 mL of BHI medium and then inoculated 1:100 into 200 µL of BHI medium containing 500 µM DOPAC and either sodium tungstate (0.5 mM, in water), sodium molybdate (0.5 mM, in water), or vehicle (water). Cultures were grown for 48 hours anaerobically at 37 °C and were harvested by centrifugation. Supernatants were dissolved 1:10 in LC-MS grade methanol and analyzed using LC-MS/MS Method F as described above. Experiments were performed anaerobically, and cultures were grown in 96-well plates (VWR, catalog# 29442-054).

### 4.4.24 Lysate assays for transcriptional and biochemical specificity of dehydroxylases from *G. pamelaeae* 3C and *E. lenta* A2

**Assays in *E. lenta* A2:** Bacterial cultures were grown in Hungate tubes. All bacterial growth and lysate experiments were performed in an anaerobic chamber. Lysis and sample processing took place in an anaerobic chamber kept at 4 °C. *E. lenta* A2 was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 50 mL of BHI medium containing 1% arginine and 10 mM formate and either 1 mM dopamine, 1 mM hydrocaffeic acid, 1 mM (+)-catechin, or vehicle. After 18 hours of anaerobic growth at 37 °C, cultures had reached OD₆₀₀ of 0.700 and were harvested by centrifugation. The bacterial pellets were resuspended anaerobically in 10 mL of cold, pre-reduced PBS to wash the cells, followed by an additional round of centrifugation to pellet the washed cells. The washed cells from each culture were then transferred to an Eppendorf tube and resuspended in 1.4 mL of lysis buffer (20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail). The cells were lysed using sonication in an anaerobic chamber. 50 µL
of this lysate was transferred in triplicate to a 96 well plate (VWR, catalog# 82006-636). 1 µL of substrate was then added to each of the replicates at a final concentration of 0.5 mM. These samples were incubated anaerobically at room temperature for 28 hours to allow for metabolism to proceed. Samples were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method B for samples containing dopamine, Method E for samples containing catechin and Method F for samples containing hydrocaffeic acid.

**Assays in Gordonibacter pamelaeae 3C:** Bacterial cultures were grown in Hungate tubes. All bacterial growth and lysate experiments were performed in an anaerobic chamber. Lysis and sample processing took place in an anaerobic chamber kept at 4 ºC. G. pamelaeae 3C was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 ºC to provide turbid starter cultures. These were diluted 1:100 in triplicate into 50 mL of BHI medium containing 10 mM formate and 1 mM DOPAC or 0.5 mM hydrocaffeic acid or vehicle. After 18 hours of anaerobic growth at 37 ºC, cultures had reached OD$_{600}$ of 0.180 and were harvested by centrifugation. The bacterial pellets were resuspended anaerobically in 10 mL of cold, pre-reduced PBS to wash the cells, followed by an additional round of centrifugation to pellet the washed cells. The washed cells from each culture were then transferred to an Eppendorf tube and resuspended in 1.4 mL of lysis buffer (20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail). The cells were lysed using sonication in an anaerobic chamber. 50 µL of this lysate was transferred in triplicate to a 96 well plate (VWR, catalog# 82006-636). 1 µL substrate was then added to each of the replicates, for a final concentration of 0.5 mM. These
samples were left anaerobically at room temperature for 28 hours to allow for metabolism to proceed. Samples were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using LC-MS/MS Method F described above.

4.4.25 Comparative genomics among human gut Actinobacteria

To characterize the distribution of Cadh, Hcdh, Dodh among our gut Actinobacterial strain library, we performed a tBLASTn search. We queried the genomes for Cadh, Hcdh, Dodh and used 90% coverage, 75% amino acid identity, and e-value=0 as the cutoff for assessing the presence of each dehydroxylase.

4.4.26 Anaerobic activity-based purification of E. lenta A2 hydrocaffeic acid dehydroxylase

All procedures were carried out under strictly anaerobic conditions at 4 °C. Procedures outside the anaerobic chamber were performed in tightly sealed containers to prevent oxygen contamination. First, E. lenta A2 starter cultures were inoculated from single colonies into liquid BHI medium and were grown for 30 hours. Starter cultures were diluted 1:100 into 8 L of BHI medium containing 1% arginine and 10 mM formate and grown anaerobically at 37 °C for 17 hours. Hydrocaffeic acid (1M stock solution in water) was added to a final concentration of 0.5 mM in the cultures. Cells were pelleted in 8 separate 1 L bottles by centrifugation (6000 rpm, 15 mins), and each pellet was resuspended in 12 mL of 20 mM Tris pH 8 containing 4 mg/mL SIGMAFAST protease inhibitor cocktail. Resuspended cells were then lysed using two rounds of sonication in an anaerobic chamber (Branson Sonifier 450, 2 min total, 10 sec on, 40 sec off, 25% amplitude). The lysates were then clarified by centrifugation (10800 rpm, 15 mins), and the soluble
fractions were subjected to two rounds of ammonium sulfate precipitation. During the precipitation, two different tubes each containing 40 mL total clarified lysate were precipitated in parallel. Solid ammonium sulfate was first dissolved in these clarified lysates to a final concentration of 30\% (w/v), and lysates were left for 1 hour and 30 minutes followed by centrifugation to pellet the precipitates (3300 g, 15 mins). The supernatant was saved, and the pellet was discarded. The supernatant was mixed with additional solid ammonium sulfate to achieve a final concentration of 40\% (w/v) and left for 1 hour and 30 minutes. Following centrifugation (3300 g, 15 mins) and removal of supernatant, each pellet containing the precipitated proteins was re-dissolved in 20 mL 20 mM Tris pH 8 containing 0.5 M ammonium sulfate. The re-dissolved pellets were combined and centrifuged to remove particulates (10800 rpm, 15 mins). The resulting 80 mL solution was injected onto an FPLC (Bio-Rad BioLogic DuoFlow System equipped with GE Life Sciences DynaLoop90) for hydrophobic interaction chromatography (HIC) using 4 x 1mL HiTrap Phenyl HP columns strung together (GE Life Sciences, catalog# 17135101). Fractions were eluted with a gradient of 0.5 M to 0 M ammonium sulfate (in 20 mM Tris pH 8) at a flow rate of 1 mL/min and were tested for activity using the assay described below. The majority of the hydrocaffeic acid dehydroxylase activity eluted between 0.2-0.5 M ammonium sulfate. Active fractions displaying >50\% conversion of hydrocaffeic acid were combined and injected onto the FPLC system described above for anion exchange chromatography using a UNO Q1 column (Bio-Rad, catalog# 720-0001) at a flow-rate of 1 mL/min. Fractions were eluted using a gradient of 0 to 1 M NaCl in 20 mM Tris pH 8 and were tested for activity. The majority of the hydrocaffeic acid dehydroxylase activity eluted between 200-400 mM NaCl. Active fractions were combined and concentrated 20-fold using a spin concentrator with a 5 kDa cutoff (3300 g centrifugation speed). 500 \mu\text{L} of the concentrate
was injected onto FPLC for size exclusion chromatography using an Enrich 24 mL column (Enrich SEC 650, 10*300 column, Bio-Rad, catalog# 780-1650). Fractions were eluted over a 26 mL volume run isocratically in 20 mM Tris pH 8 containing 250 mM NaCl and were subjected to activity assays. Active fractions were then combined and used for enzyme assays and were run on SDS-PAGE to assess the presence of protein.

**Activity assays during protein purification:** 100 µL aliquots of fractions from FPLC runs were mixed, in the following order, with 1 µL of methyl viologen (0.5 mM final concentration), 2 µL sodium dithionite (1 mM final concentration, dissolved in water), and 1 µL substrate (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for 12–14 hours to allow dehydroxylation to proceed, followed by assessment of activity using the colorimetric assay for catechol detection.

**4.4.27 Proteomics of E. lenta A2 hydrocaffeic acid dehydroxylase**

This work was performed by the Bauer Core at Harvard University. **Sample preparation:** the cut-out gel band was washed twice with 50% aqueous acetonitrile for 5 min followed by drying in a SpeedVac. The gel was then reduced with a volume sufficient to completely cover the gel pieces (100 µL) of 20 mM TCEP in 25 mM TEAB at 37 °C for 45 minutes. After cooling to room temp, the TCEP solution was removed and replaced with the same volume of 10 mM iodoacetamide Ultra (Sigma) in 25 mM TEAB and kept in the dark at room temperature for 45 minutes. Gel pieces were washed with 200 µL of 100 mM TEAB (10 minutes). The gel pieces were then shrunk with acetonitrile. The liquid was then removed followed by swelling with the 100 mM TEAB again and dehydration/shrinking with the same volume of acetonitrile. All of the liquid was removed, and the gel was completely dried in a SpeedVac for ~20 minutes. 0.06 µg/5 µL of trypsinin 50 mM TEAB was added to the gel pieces and the mixture was placed in a
thermomixer at 37°C for about 15 min. 50 µL of 50 mM TEAB was added to the gel slices. The samples were vortexed, centrifuged, and placed back in the thermomixer overnight. Samples were digested overnight at 37 °C. Peptides were extracted with 50 µL 20 mM TEAB for 20 min and 1 change of 50 µL 5% formic acid in 50% acetonitrile at room temp for 20 minutes while in a sonicator. All extracts obtained were pooled into an HPLC vial and were dried using a SpeedVac to the desired volume (~50 µL). This sample was used for protein identification by LC-MS/MS, as described below.

**Mass spectrometry:** Each sample was submitted for a single LC-MS/MS experiment that was performed on an LTQ Orbitrap Elite (Thermo Fischer) equipped with a Waters (Milford, MA) NanoAcquity HPLC pump. Peptides were separated using a 100 µm inner diameter microcapillary trapping column packed first with approximately 5 cm of C18 Reprosil resin (5 µm, 100 Å, Dr. Maisch GmbH, Germany) followed by ~20 cm of Reprosil resin (1.8 µm, 200 Å, Dr. Maisch GmbH, Germany). Separation was achieved through applying a gradient of 5–27% ACN in 0.1% formic acid over 90 min at 200 nL min⁻¹. Electrospray ionization was enabled through applying a voltage of 1.8 kV using a home-made electrode junction at the end of the microcapillary column and sprayed from fused silica pico tips (New Objective, MA). The LTQ Orbitrap Elite was operated in data-dependent mode for the mass spectrometry methods. The mass spectrometry survey scan was performed in the Orbitrap in the range of 395–1,800 m/z at a resolution of $6 \times 10^4$, followed by the selection of the twenty most intense ions (TOP20) for CID-MS2 fragmentation in the Ion trap using a precursor isolation width window of 2 m/z, AGC setting of 10,000, and a maximum ion accumulation of 200 ms. Singly charged ion species were not subjected to CID fragmentation. Normalized collision energy was set to 35 V and an activation time of 10 ms. Ions in a 10 ppm m/z window around ions selected for MS2 were excluded from
further selection for fragmentation for 60 seconds. The same TOP20 ions were subjected to HCD MS2 event in Orbitrap part of the instrument. The fragment ion isolation width was set to 0.7 m/z, AGC was set to 50,000, the maximum ion time was 200 ms, normalized collision energy was set to 27 V and an activation time of 1 ms for each HCD MS2 scan.

**Mass spectrometry data analysis:** Raw data were submitted for analysis in Proteome Discoverer 2.1.0.81 (Thermo Scientific) software. Assignment of MS/MS spectra were performed using the Sequest HT algorithm by searching the data against a protein sequence database including a custom database from *Eggerthella lenta* A2 and other sequences such as human keratins and common lab contaminants. Sequest HT searches were performed using a 20 ppm precursor ion tolerance and requiring each peptide’s N-/C-termini to adhere with trypsin protease specificity, while allowing up to two missed cleavages. Cysteine carbamidomethyl (+57.021) was set as a static modification while methionine oxidation (+15.99492 Da) was set as a variable modification. A MS2 spectra assignment false discovery rate (FDR) of 1% on protein level was achieved by applying the target-decoy database search. Filtering was performed using a Percolator (64bit version). For quantification, a 0.02 m/z window centered on the theoretical m/z value of each the six reporter ions and the intensity of the signal closest to the theoretical m/z value was recorded. Reporter ion intensities were exported in result file of Proteome Discoverer 2.1 search engine as an excel tables.

### 4.4.28 Assays of partially purified *E. lenta A2* hydrocaffeic acid dehydroxylase substrate scope

Active fractions from the size exclusion chromatography step described above were combined and then diluted 1:5 in 20 mM Tris pH 8 containing 250 mM NaCl. The enzyme mixture was transferred to the wells of a 96 well plate, for a final volume of 50 µL in each well (VWR,
1 µL of substrate (in water for hydrocaffeic acid, dopamine, and DOPAC, or 50:50 water:DMF for (+)-catechin, caffeic acid, and ferulic acid) was then added at a final concentration of 500 µM. Following this, 1 µL of a solution containing methyl viologen (1 mM final concentration) and 2 µL of sodium dithionite (2 mM final concentration, dissolved in water) were added. The resulting solution was mixed by pipetting and the 96-well plate was then sealed tightly with an aluminum seal. The enzyme assay mixtures were left at room temperature in an anaerobic chamber for 26 hours to allow dehydroxylation to proceed. The enzyme reaction mixtures were quenched by bringing the samples out of the anaerobic chamber and freezing at –20 ºC. These mixtures were then diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS using Method F for samples containing hydrocaffeic acid and DOPAC, Method E for samples containing catechin, Method C for samples containing dopamine, Method Q for samples containing caffeic acid, and Method D for samples containing ferulic acid.

4.4.29 Metagenomic analysis of catechol dehydroxylase abundance and prevalence across human patients

To generate estimates of catechol dehydroxylase enzyme prevalence in human populations, the amino acid sequences of Cldh, Dodh, Dadh, Hcdh, and Cadh were searched against a non-redundant gut microbiome gene catalogue using BLASTP with a minimum 70% percent identity, query coverage and target coverage and used to extract per-sample gene abundances from a collection of human metagenomes (10.1093/bioinformatics/btv382). High identity matches were obtained for all queries (81.7%, 81.7%, 100%, 99.5%, and 99.9% respectively over >99% target coverage) with a second lower-identity match observed for Hcdh (75.7%) for which abundances were summed with the higher identity hit. Next, to account for repeated sampling of individuals, the median gene abundance across a subject’s samples was calculated and carried forward for
prevalence estimates leading to a total of 1872 human subjects considered (58). Prevalence was then calculated as a rolling function of minimum abundance. This analysis was performed by Dr. Jordan Bisanz (Turnbaugh lab, UCSF).

**4.4.30 Incubation of human fecal samples with \( d_4 \)-dopamine, hydrocaffeic acid, (+)-catechin, and \( d_5 \)-DOPAC**

20 µL of fecal slurries from n=12 unrelated humans were diluted 1:50 into three independent cultures containing 980 µL of BHI medium and 0.5 mM substrate. As positive controls, saturated cultures of 20 µL \( E. \) lenta A2 or \( G. \) pamelaeae 3C (grown from a glycerol stock for 48 hours in BHI) were added in triplicate to 980 µL BHI medium and 0.5 mM hydrocaffeic acid, (+)-catechin, \( d_4 \)-dopamine for \( E. \) lenta A2, and \( d_5 \)-DOPAC for \( G. \) pamelaeae 3C. These cultures were grown anaerobically for 72 hours at 37 ºC. The cultures were harvested by centrifugation and 20 µL of the supernatant was then diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS using Method F for samples grown with hydrocaffeic acid, Method P for samples grown with \( d_4 \)-dopamine, Method O for samples grown with \( d_5 \)-DOPAC, and Method E for samples grown with catechin. In addition, the total community gDNA was extracted from the remaining 980 µL of culture growth with \( d_5 \)-DOPAC for downstream qPCR assays as detailed below. The experiment was performed in a 96-well plate (Agilent, catalog# A696001000).

**4.4.31 qPCR assays for dodh abundance in human fecal samples grown in BHI medium with \( d_5 \)-DOPAC**

gDNA was extracted from the culture pellets generated in the experiments described above (“Incubation of human fecal samples with \( d_4 \)-dopamine, hydrocaffeic acid, (+)-catechin, and \( d_5 \)-DOPAC”) using the DNeasy UltraClean Microbial Kit. 40 ng of the extracted DNA (2 µL) from each culture was used for qPCR assays containing 10 µL of iTaq Universal SYBRgreen Supermix
(Bio-rad, catalog 3: 1725121), 7 µL of water, and 10 µM each of forward and reverse primers. PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 95 ºC for 5 minutes 34 cycles of 95 ºC for 1 min, 60 ºC for 1 min, 72 ºC for 1 min. The program ended with a final extension at 34 ºC for 5 mins. The primers used for dodh were 5’ TACGCCTACAAACAGCTCCAA 3’ and 5’ ACATCATCTGGGGCGGATAC 3’. Data analysis was performed in GraphPad Prism (version 8).

4.4.32 Phylogenetic analysis of relationship between catechol dehydroxylases and other characterized members of the bis-MGD enzyme family

For phylogenetic analysis of the bis-MGD enzymes, we gathered protein sequences that had been previously used to study the evolution of bis-MGD enzymes (62). However, we also added sequences to capture additional diversity of biochemically characterized bis-MGD enzymes that were not included in the original tree described in (62). In particular, we performed a pBLAST search in Uniprot using perchlorate reductase (Uniprot ID# PCRA_DECAR), ethylbenzene dehydrogenase (Uniprot ID# Q5NZV2_AROAE), acetylene hydratase (Uniprot ID# AHY_PELAE), and pyrogallol transhydroxylase (Uniprot ID# PGTL_PELAC) as the queries, and collected sequences with 85-90% amino acid ID. In addition, we added the sequences of Dadh, Hcdh, Cadh from E. lenta A2, and Dodh and Cldh (19) from G. pamelaeae 3C. The sequences were combined with those reported in (62) and were aligned in Geneious (version 11) using MUSCLE. We subsequently used FastTree (standard settings, 20 rate categories of sites) to create a maximum likelihood tree. The tree files were uploaded to the Interactive Tree of Life web server (https://itol.embl.de/) to annotate the trees (109).
4.4.33 Construction of a protein sequence similarity network of the bis-MGD enzyme family.

A SSN was generated using the EFI-EST tool (http://efi.igb.illinois.edu/efi-est/) on July 15 2017 (110). In particular, we generated an SSN of the bis-MGD binding domain enzyme superfamily (PF01568), including sequences between 600 and 1400 amino acids in length and using an initial alignment score of e-150. Nodes represented sequences with 75% amino acid identity. The SSN was imported into Cytoscope v 3.2.1 and visualized with the ‘Organic layout’ setting. The alignment score cutoff was increased to e-167 until the groups of biochemically characterized enzymes included in our phylogenetic analysis separated from each other into putatively isofunctional clusters.

4.4.34 Gene neighborhood network analysis of E. lenta and Gordonibacter-type dehydroxylases

We subjected the sequences within the two separate clusters within the bis-MGD enzyme SSN harboring the E. lenta dehydroxylases (Dadh, Hcdh, Cadh) and Gordonibacter dehydroxylases (Cldh, Dodh) to gene neighborhood analysis using the EFI-EST online tool (64). We used a neighborhood size of 10 and co-occurrence frequency of 20 %. The network output was imported into Cytoscope v 3.2.1 and visualized with the ‘Organic layout’ setting. The genetic architecture output was visualized using the GNN online tool and served as the basis for the generation of a model for an average operonic architecture.

4.4.35 Phylogenetic analysis of catechol dehydroxylases encoded by gut Actinobacteria and environmental isolates

To identify additional diversity beyond the newly identified putative dehydroxylases from this study, we created a database containing putative homologs from a collection of 26 previously sequenced Actinobacterial genomes (46), as well as from genomes publicly available through
NCBI. First, the *Eggerthella lenta* A2 dopamine dehydroxylase (Dadh) protein sequence was used as the query sequence for a tBLASTn search of 26 previously sequenced Actinobacterial genomes (46) (April 23, 2019). The genomes were loaded in Geneious (version 11) and hits with an amino acid ID of >30% and e-value of e-34 were considered potential dehydroxylase hits and were saved. This cutoff was chosen because sequences captured within this window more closely resembled the acetylene hydratase and Dadh than any other biochemically characterized bis-MGD enzyme, as assessed by percent amino acid identity. In addition, we used the representative *Gordonibacter* enzyme Cldh as a separate query to identify the more distantly related, smaller enzymes from *Gordonibacter* that were not detected when using the large, multi-subunit Dadh as the query. Specifically, we used tBLASTn to search the 26 Actinobacterial genomes for the *Gordonibacter pamelaeae* 3C Cldh protein sequence. Hits from *Paraeggerthella hongongensis*, *Gordonibacter pamelaeae* 3C and *Gordonibacter sp.* 28C, the only organisms containing these smaller Cldh-like enzymes in our collection, were saved. Again, amino acid ID of >30% and e-value of e-34 were considered potential hits because sequences captured within this window more closely resembled the acetylene hydratase and Dadh than any other biochemically characterized bis-MGD enzyme, as assessed by percent amino acid identity. The hits from our searches with Cldh and Dadh were combined into a preliminary database in Geneious. To expand the sequence diversity within this database, we used Cldh and Dadh as queries for two separate tBLASTn searches in NCBI (nucleotide collection). To ensure that we captured diversity beyond human gut microbes, we excluded *Gordonibacter* and *Eggerthella* as organisms in the tBLASTn searches for Cldh and Dadh queries, respectively. For the two searches, sequences of >29% amino acid ID and e-value of e-55 were considered potential dehydroxylase hits. This was a more conservative cutoff than we used with human Actinobacteria and was selected based on the observation that the pBLAST
alignment of Dadh and Cldh has an e value of e-45 and 29% amino acid ID. The sequences retrieved from NCBI were added to the database already containing the hits from searches of the 26 Actinobacterial genomes. In addition, we added the biochemically characterized *E. coli* bis-MGD enzyme DMSO reductase (DmsA, Uniprot ID#P18775) to this database as the outgroup. This sequence was also used as the root of the tree. For phylogenetic analysis of these sequences, we first aligned sequences in Geneious using MUSCLE and removed sequences that were 95% identical to each other (considered duplicates). After deleting these duplicate sequences, we re-aligned the sequences using MUSCLE (standard settings) and subsequently used FastTree (standard settings, 20 rate categories of sites) to create a maximum likelihood tree. The tree files were uploaded to the Interactive Tree of Life web server (https://itol.embl.de/) to annotate the trees (109).

4.4.36 Phylogenetic analysis of relationship between representative dehydroxylase homologs and other characterized members of the bis-MGD enzyme family

Once we had constructed the two trees described above and uncovered dehydroxylase homologs in gut and environmental bacteria, we wanted to explore the phylogenetic relationship between these enzymes and the broader bis-molybdopterin guanine dinucleotide (bis-MGD) enzyme family. To do this, we added the representative sequences selected from the catechol dehydroxylase homolog tree to the sequence database already described in “Phylogenetic analysis of relationship between catechol dehydroxylases and other characterized members of the bis-molybdopterin guanine dinucleotide enzyme family”. Using MUSCLE, we aligned the newly added sequences with the sequences represented on the tree. We then used FastTree (standard settings, 20 rate categories of sites) in Geneious (version 11) to generate the maximum-likelihood tree. The tree files were uploaded to the Interactive Tree of Life web server
(https://itol.embl.de/) to annotate the trees (109).

### 4.4.37 Mammalian fecal samples used in this study

The collection of fecal samples from mammals (n=12 different species, n=3 individuals per species) has been previously described (81). The samples were generously provided by Dr. Aspen Reese, a Junior Fellow working with Dr. Rachel Carmody at Harvard University. To prepare these samples for culturing, all samples were resuspended anaerobically in pre-reduced PBS at a final concentration of 0.1 g/mL. The mixture was vortexed to produce a homogenous slurry and was then left for 30 minutes to let particulates settle. Aliquots of the supernatant were dissolved 50:50 with 40% glycerol in water and flash-frozen in liquid nitrogen, creating slurries. These slurries were stored at –80 ºC and were thawed anaerobically at room temperature at the time of use.

### 4.4.38 Screen for catechol dehydroxylation by mammalian gut microbiota samples

Mammalian fecal slurries were prepared as described above and were thawed by incubation at room temperature at the time of use. 20 µL of each slurry was then combined with 980 µL of basal medium containing 500 µM each of dopamine, (+)-catechin, DOPAC, or hydrocaffeic acid. Each individual sample was grown in one well, with the n=3 individual samples for each animal serving as the biological replicates. Control wells contained compound but no bacteria. Samples were grown anaerobically at 37 ºC for 96 hours in a 96-well plate (Agilent Technology, catalog# A696001000). Following growth, we first assessed the total microbial growth by measuring the OD$_{600}$ in a plate reader (BioTek Synergy HTX). Cultures were then harvested by centrifugation and 50 µL of supernatant was transferred to a new 96-well plate, at which time the catechol colorimetric assay was used to assess total dehydroxylation by the complex microbial community. Samples that had potential catechol depletion as assessed by the colorimetric assay were then
further analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was
diluted 1:10 with 180 µL of methanol, followed by centrifugation at 4000 rpm for 10 minutes to
pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to
the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using
Method A for samples containing dopamine, Method E for samples containing catechin, and
Method F for samples containing DOPAC and hydrocaffeic acid.

4.4.39 Construction of a mammalian phylogenetic tree

The mammalian phylogenetic tree was generated using the Automatic Phylogenetic Tree
Generator (aptg, version 0.1.0) script in R (version 3.5.1). Mammals not part of the aptg database
were added manually to the tree using additional information about the mammalian phylogeny as
a reference (81). The mammalian icons were adapted under a Creative Commons license
(https://creativecommons.org/licenses/by/3.0/) at phylopic (http://phylopic.org), including Alpaca
logo (made by Steven Traver), Bison (Lukasiniho), Cow (Steven Traver), Dog (Tracy A Heath),
Fox (Anthony Caravaggi), Guinea pig (Zimices), Mouse (Madeleine Price Ball), Pig (Steven
Traver), Rabbit (Steven Traver), Rabbit (Steven Traver), Rat (Rebecca Groom), Sheep (Zimices),
and Wolf (Tracy A. Heath).
4.5 Materials and methods for synthetic chemistry

The synthetic work was performed by Dr. Michael Luescher in our laboratory, who also wrote the detailed methods section provided below.

4.5.1 General materials and methods

All reactions were performed in dried glassware under an atmosphere of dry N\textsubscript{2}. Reaction mixtures were stirred magnetically unless otherwise indicated and monitored by thin layer chromatography (TLC) on Merck precoated glass-backed silica gel 60 F-254 0.25 mm plates with visualization by fluorescence quenching at 254 nm. TLC plates were stained using a potassium permanganate solution. Chromatographic purification of products (flash column chromatography) was performed on Silicycle Silica Flash F60 (230–400 Mesh) silica gel using a forced flow of eluent at 0.3–0.5 bar. Concentration of reaction product solutions and chromatography fractions under reduced pressure was performed by rotary evaporation at 35–40 °C at the appropriate pressure and then at rt, ca. 0.1 mmHg (vacuum pump) unless otherwise indicated.

All chemicals were purchased from Acros, Aldrich, Fluka, Merck, ABCR, TCI, Alfa Aesar or Strem and used as such unless stated otherwise. Commercial grade reagents and solvents were used without further purification except as indicated below. Toluene, diethylether (Et\textsubscript{2}O), tetrahydrofuran (THF) and dichloromethane (CH\textsubscript{2}Cl\textsubscript{2}) were purified by pressure filtration through activated alumina. N,N-Dimethylformamide (DMF), acetonitrile (CH\textsubscript{3}CN), and ethanol (EtOH) were used as purchased. Yields given refer to chromatographically purified and spectroscopically pure compounds unless otherwise stated.

Infrared (IR) spectra were recorded on a Bruker ALPHA FT-IR spectrophotometer and reported as wavenumber (cm\textsuperscript{-1}) of the absorption maxima for the range between 4000 cm\textsuperscript{-1} and 750 cm\textsuperscript{-1} with only major peaks reported. ^1H NMR and ^13C NMR spectra were recorded on a Varian-
Inova-500 500 MHz, 125 MHz spectrometer. $^1$H NMR chemical shifts are expressed in parts per million (δ) downfield from tetramethylsilane (with the CHCl$_3$ peak at 7.26 ppm, MeOH peak at 3.31, DMSO peak at 2.50, and acetone peak at 2.05 used as a standard). $^{13}$C NMR chemical shifts are expressed in parts per million (δ) downfield from tetramethylsilane (with the central peak of CHCl$_3$ at 77.16 ppm, MeOH peak 49.00, DMSO peak at 39.52, and acetone peak at 29.84 used as a standard). All $^{13}$C spectra were measured with complete proton decoupling. NMR coupling constants (J) are reported in Hertz (Hz), and splitting patterns are indicated as follows: br, broad; s, singlet; d, doublet; dd, doublet of doublet; t, triplet; m, multiplet. High-resolution mass spectrometric measurements (HRMS) were performed on an Accurate-Mass 6530 Q-TOF LC/MS (Agilent) using dual electrospray ionization (ESI).
4.5.2 Preparation of 2-aminoethylbenzendiol/-triol derivatives

All reactions were carried out with degassed solvents under a positive pressure of nitrogen.

4.5.2.1 2-phenylacetic acid derivatives as starting materials

![Chemical diagram](image)

**Figure 4.52 Preparation of 2-phenylacetic acid derivatives.**

**Figure 4.53 Reduction of benzyl carboxylic acids.**

**General C’Acid Reduction Procedure:** BH$_3$ • SMe$_2$ (2.0 M in THF; 1.30 equiv) was added dropwise to a solution of commercially available di- / trimethoxyphenylacetic acid (1.00 equiv) in THF (0.25 M) at 0 °C. The resulting mixture was allowed to warm to rt over 3 h and stirring was continued for 14 h while a colorless solid formed. The obtained suspension was cooled to 0 °C and carefully quenched with the dropwise addition of saturated aqueous NaHCO$_3$. The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with H$_2$O (2 x 20 mL), brine (2 x 20 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure to yield analytically pure alcohol S1–S5 that was used in the next step without further purification.
Figure 4.54 Chemical structure of dimethoxyphenethyl alcohol (S1).

2-(2,3-Dimethoxyphenyl)ethan-1-ol (S1). Following the general carboxylic acid reduction procedure using 2-(2,3-dimethoxyphenyl)acetic acid (2.00 g, 10.2 mmol), alcohol S1 was obtained as a colorless oil (1.85 g, quant.). $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 7.01 (t, $J = 7.9$ Hz, 1H), 6.88 – 6.75 (m, 2H), 3.87 (s, 3H), 3.84 (s, 3H), 3.84 (t, $J = 6.5$ Hz, 2H), 2.91 (t, $J = 6.5$ Hz, 2H), 1.76 (br s, OH); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 152.8, 147.4, 132.6, 124.2, 122.6, 111.0, 63.4, 60.7, 55.7, 33.8. The spectral characteristics were identical to those reported in the current literature, which fails to report the signals for the OMe and one of the CH$_2$ groups (111).

Figure 4.55 Chemical structure of dimethoxyphenethyl alcohol (S2).

2-(3,5-Dimethoxyphenyl)ethan-1-ol (S2). Following the general carboxylic acid reduction procedure using 2-(3,5-dimethoxyphenyl)acetic acid (1.00 g, 5.10 mmol), alcohol S2 was obtained as a colorless oil (920 mg, quant.). $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 6.39 (d, $J = 2.4$ Hz, 2H), 6.36 – 6.33 (m, 1H), 3.86 (t, $J = 6.4$ Hz, 2H), 3.79 (s, 6H), 2.82 (t, $J = 6.4$ Hz, 2H), 1.45 (br s, OH). The spectral characteristics were identical to those reported in the current literature (112).
Figure 4.56 Chemical structure of dimethoxyphenethyl alcohol (S3).

2-(2,5-Dimethoxyphenyl)ethan-1-ol (S3). Following the general carboxylic acid reduction procedure using 2-(2,5-dimethoxyphenyl)acetic acid (2.00 g, 10.2 mmol), alcohol S3 was obtained as a colorless oil (1.70 g, 92% yield). $^1$H NMR (500 MHz, CDCl$_3$): δ 6.80 (d, $J$ = 8.6 Hz, 1H), 6.77 – 6.71 (m, 2H), 3.83 (t, $J$ = 6.4 Hz, 2H), 3.79 (s, 3H), 3.76 (s, 3H), 2.88 (t, $J$ = 6.4 Hz, 2H), 1.72 (br s, OH). The spectral characteristics were identical to those reported in the current literature (113).

Figure 4.57 Chemical structure of trimethoxyphenethyl alcohol (S4).

2-(2,3,4-Trimethoxyphenyl)ethan-1-ol (S4). Following the general carboxylic acid reduction procedure using 2-(2,3,4-trimethoxyphenyl)acetic acid (2.00 g, 8.84 mmol), alcohol S4 was obtained as a colorless oil (1.87 g, quant.). IR (thin film) ν 3348, 2926, 2850, 1769, 1658, 1602, 1498, 1395, 1091, 840 cm$^{-1}$; $^1$H NMR (500 MHz, CDCl$_3$): δ 6.86 (d, $J$ = 8.4 Hz, 1H), 6.63 (d, $J$ = 8.4 Hz, 1H), 3.90 (s, 3H), 3.87 (s, 3H), 3.84 (s, 3H), 3.80 (t, $J$ = 6.4 Hz, 2H), 2.83 (t, $J$ = 6.4 Hz, 2H), 1.79 (br s, OH); $^{13}$C NMR (125 MHz, CDCl$_3$): δ 152.6, 152.1, 142.4, 124.7, 124.7, 107.5, 63.5, 61.0, 60.8, 56.1, 33.6; ESI-HRMS calcd for C$_{11}$H$_{17}$O$_4$ [M+H] 213.1121, found 213.1128.
**Figure 4.58** Chemical structure of trimethoxyphenethyl alcohol (S5).

2-(2,4,6-Trimethoxyphenyl)ethan-1-ol (S5). Following the general carboxylic acid reduction procedure using 2-(2,4,6-trimethoxyphenyl)acetic acid (1.00 g, 4.42 mmol), alcohol S5 was obtained as a colorless oil (940 mg, quant.). $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 6.14 (s, 2H), 3.80 (s, 3H), 3.80 (s, 6H), 3.71 (t, $J = 6.4$ Hz, 3H), 2.89 (t, $J = 6.4$ Hz, 2H), 1.96 (br s, OH); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 159.8, 159.2, 107.6, 90.7, 63.1, 55.8, 55.5, 26.2. The spectral characteristics were identical to those reported in the current literature (114).

**Figure 4.59** Scheme for Mitsunobu reaction on phenethyl alcohols.

**General Mitsunobu Reaction Protocol:** Diethyl azodicarboxylate (DEAD; 40% in toluene; 1.10 equiv) was added dropwise over 5–10 min to a solution of PPh$_3$ (1.15 equiv), phthalimide (1.15 equiv) and the corresponding alcohol S1–S5 (1.00 equiv) in THF (0.15 M) at 0 °C. The resulting mixture was allowed to warm to rt over 3 h and stirring was continued for 14 h. The resulting pale-yellow solution was concentrated under reduced pressure and purified by flash column chromatography to afford the desired phthalimide protected amine S6–S10.

**Figure 4.60** Chemical structure of dimethoxyphenethyl phthalimide (S6).
2-(2,3-Dimethoxyphenethyl)isoindoline-1,3-dione (S6). Following the general Mitsunobu reaction protocol, purification by flash column chromatography (hexanes:EtOAc 5:1) afforded phthalimide protected amine S6 as a colorless solid (2.90 g, 85% yield) using alcohol S1 (2.00 g, 11.0 mmol) as starting material. ¹H NMR (500 MHz, CDCl₃): δ 7.85 – 7.78 (m, 2H), 7.73 – 7.66 (m, 2H), 6.94 (t, J = 7.9 Hz, 1H), 6.83 – 6.74 (m, 2H), 3.95 – 3.91 (m, 2H), 3.89 (s, 3H), 3.83 (s, 3H), 3.07 – 2.97 (m, 2H); ¹³C NMR (125 MHz, CDCl₃): δ 168.2, 152.8, 147.7, 133.8, 132.2, 131.9, 123.8, 123.1, 122.3, 111.3, 60.8, 55.7, 38.5, 29.1. The spectral characteristics were identical to those reported in the current literature (115).

![Figure 4.61 Chemical structure of Dimethoxyphenethyl phthalimide (S7).](image)

2-(3,5-Dimethoxyphenethyl)isoindoline-1,3-dione (S7). Following the general Mitsunobu reaction protocol, purification by flash column chromatography (hexanes:EtOAc 3:1) afforded phthalimide protected amine S7 as a colorless solid (1.70 g, 99% yield) using alcohol S2 (1.00 g, 5.50 mmol) as starting material. ¹H NMR (500 MHz, CDCl₃): δ 7.84 (dd, J = 5.4, 3.1 Hz, 2H), 7.71 (dd, J = 5.4, 3.1 Hz, 2H), 6.45 – 6.38 (m, 2H), 6.32 (t, J = 2.2 Hz, 1H), 3.96 – 3.89 (m, 2H), 3.75 (s, 6H), 2.96 – 2.91 (m, 2H); ¹³C NMR (125 MHz, CDCl₃): δ 168.3, 161.0, 140.4, 134.0, 132.2, 123.3, 106.8, 99.0, 55.4, 39.2, 35.0; Rᵣ = 0.33 (hexanes:EtOAc 3:1). The spectral characteristics were identical to those reported in the current literature (115).

![Figure 4.62 Chemical structure of imethoxyphenethyl phthalimide (S8).](image)
2-(2,5-Dimethoxyphenethyl)isoindoline-1,3-dione (S8). Following the general Mitsunobu reaction protocol, purification by flash column chromatography (hexanes:EtOAc 5:1) afforded phthalimide protected amine S8 as a colorless solid (3.15 g, 98% yield) using alcohol S3 (1.85 g, 10.2 mmol) as starting material. \(^1\)H NMR (500 MHz, CDCl\(_3\)): δ 7.79 – 7.71 (m, 2H), 7.68 – 7.59 (m, 2H), 6.70 – 6.63 (m, 3H), 3.90 (t, J = 7.1 Hz, 2H), 3.66 (s, 3H), 3.62 (s, 3H), 2.93 (t, J = 7.1 Hz, 2H); \(^{13}\)C NMR (125 MHz, CDCl\(_3\)): δ 168.1, 153.3, 152.0, 133.7, 132.1, 127.6, 123.0, 116.6, 112.2, 111.1, 55.7, 55.6, 37.8, 29.6; R\(_f\) = 0.28 (hexanes:EtOAc 5:1). The spectral characteristics were identical to those reported in the current literature (115).

Figure 4.63 Chemical structure of trimethoxyphenethyl phthalimide (S9).

2-(2,3,4-Trimethoxyphenethyl)isoindoline-1,3-dione (S9). Following the general Mitsunobu reaction protocol, purification by flash column chromatography (hexanes:EtOAc 5:1) afforded phthalimide protected amine S9 as a colorless solid (1.55 g, 96% yield) using alcohol S4 (1.00 g, 4.71 mmol) as starting material. \(^1\)H NMR (500 MHz, CD\(_2\)OD): δ 7.85 – 7.70 (m, 4H), 6.78 (d, J = 8.5 Hz, 1H), 6.60 (d, J = 8.5 Hz, 1H), 3.86 (t, J = 6.8 Hz, 2H), 3.82 (s, 3H), 3.77 (s, 3H), 3.65 (s, 3H), 2.93 – 2.87 (m, 2H); \(^{13}\)C NMR (125 MHz, CD\(_2\)OD): δ 169.7, 154.1, 153.5, 143.4, 135.3, 135.2, 126.0, 124.1, 123.9, 108.7, 61.3, 60.9, 56.4, 39.7, 29.8; R\(_f\) = 0.18 (hexanes:EtOAc 3:1). The spectral characteristics were identical to those reported in the current literature (115).
Figure 4.64 Chemical structure of trimethoxyphenethyl phthalimide (S10).

2-(2,4,6-Trimethoxyphenethyl)isoindoline-1,3-dione (S10). Following the general Mitsunobu reaction protocol, purification by flash column chromatography (hexanes:EtOAc 3:1) afforded phthalimide protected amine S9 as a colorless solid (1.60 g, 99% yield) using alcohol S5 (1.00 g, 4.71 mmol) as starting material. IR (thin film) ν 3248, 2928, 2850, 1772, 1715, 1634, 1607, 1439, 1205, 1139, 809 cm\(^{-1}\); \(^1\)H NMR (500 MHz, CDCl\(_3\)): δ 7.84 – 7.73 (m, 2H), 7.71 – 7.62 (m, 2H), 6.00 (s, 2H), 3.85 (t, \(J = 6.3\) Hz, 2H), 3.76 (s, 3H), 3.59 (s, 6H), 2.97 (t, \(J = 6.3\) Hz, 2H); \(^{13}\)C NMR (125 MHz, CDCl\(_3\)): δ 168.4, 160.0, 159.2, 133.6, 132.5, 122.9, 107.4, 90.2, 55.5, 55.3, 37.6, 21.6; \(R_f = 0.20\) (hexanes:EtOAc 3:1); ESI-HRMS calcd for C\(_{18}\)H\(_{18}\)NO\(_4\) [M+H] 312.1230, found 312.1223.

Figure 4.65 Hydrazinolysis of phenethyl phthalimides.

General Amine Phthalimide Deprotection Protocol: Hydrazine monohydrate (10.0 equiv) was added to a suspension of the phthalimide protected amine S6 – S10 in EtOH (0.15 M). The resulting solution was heated to reflux for 1.5 h while colorless solids crashed out. The resulting suspension was allowed to cool to rt before H\(_2\)O (20–50 mL) was added in one portion. Stirring was continued to afford a clear solution that was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with H\(_2\)O (3 x 20 mL), brine (2 x 20 mL), dried over anhydrous...
Na₂SO₄, filtered and concentrated under reduced pressure to yield analytically pure amine S11–S15 that was used in the next step without further purification.

![Chemical structure of dimethoxyphenethyl amine (S11).](image1)

**2-(2,3-Dimethoxyphenyl)ethan-1-amine (S11).** Following the general phthalimide deprotection protocol using phthalimide protected amine S6 (1.00 g, 3.21 mmol), primary amine S11 was obtained as a colorless oil (500 mg, 86% yield). ¹H NMR (500 MHz, CD₃OD): δ 6.99 (dd, J = 8.0, 7.7 Hz, 1H), 6.88 (dd, J = 8.0, 1.5 Hz, 1H), 6.78 (dd, J = 7.7, 1.6 Hz, 1H), 3.84 (s, 3H), 3.79 (s, 3H), 2.88 – 2.80 (m, 2H), 2.80 – 2.73 (m, 2H); NH₂-group is not visible; ¹³C NMR (125 MHz, CD₃OD): δ 154.2, 148.6, 134.3, 125.1, 123.4, 112.1, 61.0, 56.2, 43.5, 34.5. The spectral characteristics were identical to those reported in the current literature (116).

![Chemical structure of dimethoxyphenethyl amine (S12).](image2)

**2-(3,5-Dimethoxyphenyl)ethan-1-amine (S12).** Following the general phthalimide deprotection protocol using phthalimide protected amine S7 (400 mg, 1.28 mmol), primary amine S12 was obtained as a colorless oil (140 mg, 60% yield). ¹H NMR (500 MHz, CD₃OD): δ 6.41 – 6.36 (m, 2H), 6.35 (t, J = 2.3 Hz, 1H), 3.76 (s, 6H), 2.92 (t, J = 7.3 Hz, 2H), 2.72 (t, J = 7.3 Hz, 2H), NH₂-group is not visible; ¹³C NMR (125 MHz, CD₃OD): δ 162.5, 142.6, 107.8, 99.3, 55.7, 43.6, 39.4. The spectral characteristics were identical to those reported in the current literature (117).
Figure 4.68 Chemical structure of dimethoxyphenethyl amine (S13).

2-(2,5-Dimethoxyphenyl)ethan-1-amine (S13). Following the general phthalimide deprotection protocol using phthalimide protected amine S8 (500 mg, 1.61 mmol), primary amine S13 was obtained as a colorless oil (250 mg, 86% yield). $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 6.96 – 6.57 (m, 3H), 3.78 (s, 3H), 3.76 (s, 3H), 3.04 – 2.85 (m, 3H), 2.75 (t, $J$ = 6.9 Hz, 2H), 1.90 (br s, NH$_2$); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 153.5, 152.0, 129.2, 117.0, 111.4, 111.3, 55.9, 55.7, 42.1, 34.6. The spectral characteristics were identical to those reported in the current literature ($^{118}$).

Figure 4.69 Chemical structure of trimethoxyphenethyl amine (S14).

2-(2,3,4-Trimethoxyphenyl)ethan-1-amine (S14). Following the general phthalimide deprotection protocol using phthalimide protected amine S9 (1.40 g, 4.10 mmol), primary amine S14 was obtained as a colorless oil (600 mg, 69% yield). $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 6.83 (d, $J$ = 8.5 Hz, 1H), 6.61 (d, $J$ = 8.5 Hz, 1H), 3.87 (s, 3H), 3.86 (s, 3H), 3.83 (s, 3H), 2.90 (t, $J$ = 7.0 Hz, 2H), 2.70 (t, $J$ = 7.0 Hz, 2H), 1.51 (br s, NH$_2$); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 152.4, 152.2, 142.5, 125.8, 124.5, 107.3, 61.1, 60.8, 56.1, 43.1, 34.2. The spectral characteristics were identical to those reported in the current literature ($^{116}$).

Figure 4.70 Chemical structure of trimethoxyphenethyl amine (S15).
**2-(2,4,6-Trimethoxyphenyl)ethan-1-amine (S15).** Following the general phthalimide deprotection protocol using phthalimide protected amine S10 (1.60 g, 4.69 mmol), primary amine S15 was obtained as a colorless oil (693 mg, 70% yield). IR (thin film) ν 3426, 2938, 2838, 1593, 1498, 1455, 1417, 1204, 1148 cm⁻¹; ¹H NMR (500 MHz, CD₃OD): δ 6.20 (s, 2H), 3.79 (s, 6H), 3.79 (s, 3H), 2.79 – 2.71 (m, 4H), NH₂-group is not visible; ¹³C NMR (125 MHz, CD₃OD): δ 161.5, 160.4, 108.1, 91.5, 56.0, 56.0, 55.7, 41.9, 25.8; ESI-HRMS calcd for C₁₁H₁₈NO₃ [M+H] 212.1281, found 212.1280.

![Figure 4.71 Practical demethylation of aryl methyl ethers.](image)

**General Phenol Ether Cleavage Protocol:** BBr₃ (1.0 M in CH₂Cl₂; 1.15 equiv for each OMe group) was added dropwise to a solution of the 2-(methoxyphenyl)ethyl amine S11–S15 (1.00 equiv) in CH₂Cl₂ (0.025 M) at –78 °C. The resulting mixture was allowed to warm to rt over 3 h. Stirring was continued for 14 h. The resulting suspension was cooled to 0 °C and quenched with the dropwise addition of MeOH (ca. 5 mL). Stirring at rt was continued for 1 h. The resulting suspension was concentrated under reduced pressure to afford a pale-brown oil. The obtained residue was dissolved in a small amount of MeOH and again concentrated under reduced pressure; this step was repeated 3–4 times to remove all of the trimethyl borate side product and obtain analytically pure 2-(2-aminoethyl)benzenediol /-triol derivatives S16–S20 as HBr salts.

![Figure 4.72 Chemical structure of aminoethylbenzene diol hydrobromide (S16).](image)
3-(2-Aminoethyl)benzene-1,2-diol hydrobromide (S16). Following the general phenol ether cleavage protocol using phenol ether S11 (200 mg, 1.10 mmol), catechol amine S16 was obtained as a brown oil (250 mg, 97% yield). $^1$H NMR (500 MHz, CD$_3$OD): $\delta$ 6.73 (dd, $J = 6.7$, 2.7 Hz, 1H), 6.69 – 6.56 (m, 2H), 3.17 (t, $J = 7.4$ Hz, 2H), 2.95 (t, $J = 7.4$ Hz, 2H), OH- and NH-protons are not visible; $^{13}$C NMR (125 MHz, CD$_3$OD): $\delta$ 146.0, 144.6, 124.4, 122.3, 120.8, 115.4, 40.8, 29.4. The spectral characteristics were identical to those reported in the current literature (119).

Figure 4.73 Chemical structure of aminoethylbenzene diol hydrobromide (S17).

5-(2-Aminoethyl)benzene-1,3-diol hydrobromide (S17). Following the general phenol ether cleavage protocol using phenol ether S12 (20 mg, 0.110 mmol), resorcinol amine S17 was obtained as a brown oil (25.0 mg, 97% yield). IR (thin film) $\nu$ 3358, 2928, 2853, 1771, 1597, 1495, 1418, 1091, 844 cm$^{-1}$; $^1$H NMR (500 MHz, CD$_3$OD): $\delta$ 6.33 – 6.08 (m, 2H), 3.13 (t, $J = 7.6$ Hz, 2H), 2.80 (t, $J = 7.6$ Hz, 2H), OH- and NH-protons are not visible; $^{13}$C NMR (125 MHz, CD$_3$OD): $\delta$ 160.0, 139.9, 108.1, 102.4, 41.9, 34.5; ESI-HRMS calcd for C$_8$H$_{12}$NO$_2$ [M+H] 154.0863, found 154.0860.

Figure 4.74 Chemical structure of 5-(2-Aminoethyl)benzene-1,4-diol hydrobromide (S18).
**5-(2-Aminoethyl)benzene-1,4-diol hydrobromide (S18).** Following the general phenol ether cleavage protocol using phenol ether S13 (100 mg, 0.552 mmol), diol amine S18 was obtained as a brown oil (118 mg, 91% yield). IR (thin film) \( \nu \) 3352, 2927, 2858, 1621, 1505, 1455, 1344, 1202, 1152, 1212 cm\(^{-1}\); \(^1\)H NMR (500 MHz, CD\(_3\)OD): \( \delta \) 6.66 (d, \( J = 8.5 \) Hz, 1H), 6.63 – 6.46 (m, 2H), 3.15 (t, \( J = 7.2 \) Hz, 2H), 2.88 (t, \( J = 7.2 \) Hz, 2H), OH- and NH-protons are not visible; \(^{13}\)C NMR (125 MHz, CD\(_3\)OD): \( \delta \) 151.1, 149.4, 124.9, 118.2, 116.9, 115.8, 40.9, 29.9; ESI-HRMS calcd for C\(_8\)H\(_{12}\)NO\(_2\) [M+H] 154.0863, found 154.0852.

![Figure 4.75 Chemical structure of aminoethylbenzene triol hydrobromide (S19).](image)

**4-(2-Aminoethyl)benzene-1,2,3-triol hydrobromide (S19).** Following the general phenol ether cleavage protocol using phenol ether S14 (200 mg, 0.948 mmol), triol amine S19 was obtained as a brown oil (240 mg, quant.). IR (thin film) \( \nu \) 3357, 3222, 2537, 1620, 1484, 1282, 1230, 1182, 1100, 1053, 1016 cm\(^{-1}\); \(^1\)H NMR (500 MHz, CD\(_3\)OD): \( \delta \) 6.48 (d, \( J = 8.2 \) Hz, 1H), 6.32 (d, \( J = 8.2 \) Hz, 1H), 3.12 (t, \( J = 7.3 \) Hz, 2H), 2.87 (t, \( J = 7.3 \) Hz, 2H), OH- and NH-protons are not visible; \(^{13}\)C NMR (125 MHz, CD\(_3\)OD): \( \delta \) 146.6, 145.7, 134.4, 121.4, 115.9, 108.0, 41.2, 29.5; ESI-HRMS calcd for C\(_8\)H\(_{12}\)NO\(_3\) [M+H] 170.0812, found 170.0805.

![Figure 4.76 Chemical structure of aminoethylbenzene triol hydrobromide (S20).](image)
2-(2-Aminoethyl)benzene-1,3,5-triol hydrobromide (S20). Following the general phenol ether cleavage protocol using phenol ether S15, triol amine S20 was obtained, according to MS identification, in low quantities along with brominated species and various methoxybenzene-diols in an inseparable mixture. Initial attempts to alter reaction temperature or the number of equivalents of BBr₃ resulted in low conversion. Heating phenol ether S15 in the presence of iodo(trimethyl)silane (S21) or sodium ethanethiolate (S22) afforded mono-deprotected material in a cleaner reaction, but the desired triol amine S20 was not observed. Due to our inability to access this substrate, we did not evaluate this substrate in any enzyme reactions in our study.

4.5.2.2 Benzaldehydes as starting materials

![Scheme showing the synthesis of aminoethylbenzene triol hydrobromide from aryl nitriles](image)

Figure 4.77 Preparation of aminoethylbenzene triol hydrobromide from aryl nitriles.
2,3,5-Trimethoxybenzonitrile (S21). K$_2$CO$_3$ (1.90 g, 13.8 mmol, 1.50 equiv) and dimethyl sulfate (0.960 mL, 1.28 g, 10.1 mmol, 1.10 equiv) were added to a solution of 5-hydroxy-2,3-dimethoxybenzonitrile (122) (1.65 g, 9.21 mmol, 1.00 equiv) in acetone (30 mL) at rt. Stirring was continued for 18 h to afford a pale beige suspension. The solvent was removed under reduced pressure and the resulting crude material was diluted with a mixture of EtOAc–H$_2$O (1:1; 100 mL). The obtained layers were separated and the aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with 5% aqueous NaOH (20 mL) and brine (2 x 20 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure to afford analytically pure trimethoxybenzonitrile S21 (1.78 g, quant.) as a pale beige solid. $^1$H NMR (500 MHz, CDCl$_3$): δ 6.68 (d, $J$ = 2.8 Hz, 1H), 6.56 (d, $J$ = 2.8 Hz, 1H), 3.94 (s, 3H), 3.86 (s, 3H), 3.79 (s, 3H). The spectral characteristics were identical to those reported in the current literature (123).

2,3,5-Trimethoxybenzaldehyde (S22). DIBAL-H (1.0 M in CH$_2$Cl$_2$; 12.4 mL, 12.4 mmol, 1.50 equiv) was added dropwise to a solution of nitrile S21 (1.60 g, 8.28 mmol, 1.00 equiv) in CH$_2$Cl$_2$ (33 mL) at 0 °C. The resulting mixture was allowed to warm to rt over 3 h after which stirring was continued for 8 h. The reaction was cooled to 0 °C and HCl (1.0 M in H$_2$O; 10.0 mL) was added dropwise over 10 min. The mixture was allowed to warm to rt and stirring was
continued for 2 h. The layers were separated, and the aqueous layer was extracted with CH$_2$Cl$_2$ (2 x 15 mL). The combined organic layers were washed with H$_2$O (2 x 15 mL) and brine (20 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated under reduced pressure to afford analytically pure benzaldehyde S22 (845 mg, 52%) as a beige solid. $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 10.40 (s, 1H), 6.86 (d, $J$ = 2.9 Hz, 1H), 6.74 (d, $J$ = 2.9 Hz, 1H), 3.93 (s, 3H), 3.89 (s, 3H), 3.82 (s, 3H); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 190.2, 156.4, 154.4, 148.2, 129.9, 107.7, 99.7, 63.1, 56.4, 56.1. The spectral characteristics were identical to those reported in the current literature (124).

Figure 4.80 Chemical structure of nitrovinyl benzene (S23).

(E)-1,2,5-Trimethoxy-3-(2-nitrovinyl)benzene (S23). A mixture of benzaldehyde S22 (845 mg, 4.31 mmol, 1.00 equiv) and ammonium acetate (500 mg, 6.46 mmol, 1.50 equiv) in nitromethane (40 mL) was heated to reflux for 18 h, after which the reaction was found to be complete according to TLC ($R_f$ = 0.38 starting material; $R_f$ = 0.35 product; hexanes:EtOAc 4:1). The resulting mixture was concentrated under reduced pressure and purified by flash column chromatography (hexanes:EtOAc 6:1) to afford nitrovinyl benzene S23 (750 mg, 73%) as a yellow solid. IR (thin film) $\nu$ 2959, 2846, 1717, 1633, 1601, 1492, 1465, 1332, 1282, 1206, 1176, 1151 cm$^{-1}$; $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 8.16 (d, $J$ = 13.8 Hz, 1H), 7.71 (d, $J$ = 13.8 Hz, 1H), 6.61 (d, $J$ = 2.6 Hz, 1H), 6.48 (d, $J$ = 2.6 Hz, 1H), 3.86 (s, 3H), 3.83 (s, 3H), 3.80 (s, 3H); $^{13}$C NMR (125 MHz, CDCl$_3$): $\delta$ 156.4, 154.1, 144.2, 138.7, 134.8, 124.0, 104.5, 102.9, 61.6, 56.1, 55.9; $R_f$ = 0.35 (hexanes:EtOAc 4:1); ESI-HRMS calcd for C$_{11}$H$_{14}$NO$_5$ [M+H] 240.0866, found 240.0871.
2-(2,3,5-Trimethoxyphenyl)ethan-1-amine (S24). LiAlH₄ (2.0 M in THF; 1.83 mL, 3.66 mmol, 3.50 equiv) was added dropwise over 10 min to a solution of nitrovinyl benzene S23 (250 mg, 1.05 mmol, 1.00 equiv) in THF (6 mL) at 0 °C. The resulting mixture was allowed to warm to rt and stirring was continued for 24 h. The reaction mixture was then cooled to 0 °C and 10% aqueous NaOH (5.0 mL) was added dropwise over 10 min, resulting in an exothermic reaction. Stirring was continued for 1 h, and the resulting suspension was diluted with EtOAc (20 mL) and filtered over a plug of Celite® (EtOAc rinse). The filtrate was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure to afford crude amine S24. Purification by flash column chromatography (EtOAc:MeOH 85:15 + 0.1% Et₃N) afforded amine S24 (150 mg, 68%) as a pale yellow oil. IR (thin film) v 3363, 2937, 2839, 1599, 1492, 1465, 1427, 1380, 1220, 1175, 1150, 1089, 830 cm⁻¹; ¹H NMR (500 MHz, CDCl₃): δ 6.39 (d, J = 2.5 Hz, 1H), 6.29 (d, J = 2.5 Hz, 1H), 3.84 (s, 3H), 3.77 (s, 3H), 3.76 (s, 3H), 2.97 (t, J = 7.0 Hz, 2H), 2.78 (t, J = 7.0 Hz, 2H), 2.27 (br s, NH₂); ¹³C NMR (125 MHz, CDCl₃): δ 156.0, 153.5, 141.5, 133.4, 105.5, 98.5, 60.9, 55.7, 55.6, 42.7, 34.1; Rₛ = 0.08 (EtOAc:MeOH 85:15); ESI-HRMS calcd for C₁₁H₁₈NO₃ [M+H] 212.1281, found 212.1275.
3-(2-Aminoethyl)benzene-1,2,5-triol hydrobromide (S25). BBr₃ (1.0 M in CH₂Cl₂; 1.41 mL, 1.41 mmol, 3.30 equiv) was added dropwise over 10 min to a solution of phenol ether S24 (90.0 mg, 0.425 mmol, 1.00 equiv) in CH₂Cl₂ (0.033 M) at –78 °C. The resulting mixture was allowed to warm to rt over 3 h. Stirring was continued for 18 h. The resulting suspension was cooled to 0 °C and quenched with the dropwise addition of MeOH (ca. 5 mL). Stirring at rt was continued for 1 h. The resulting solution was concentrated under reduced pressure to afford a pale-brown oil. The obtained residue was dissolved in a small amount of MeOH and again concentrated under reduced pressure; this step was repeated 3–4 times to remove all of the trimethyl borate side product and obtain analytically pure triol amine S25 as the HBr salt. IR (thin film) ν 3358, 3223, 1604, 1452, 1359, 1108, 1044 cm⁻¹; ¹H NMR (500 MHz, CD₃OD): δ 6.29 (d, J = 2.6 Hz, 1H), 6.11 (d, J = 2.8 Hz, 1H), 3.15 (t, J = 7.4 Hz, 2H), 2.88 (t, J = 7.3 Hz, 2H), OH- and NH-protons are not visible; ¹³C NMR (125 MHz, CD₃OD): δ 151.4, 147.1, 137.7, 125.0, 108.2, 103.3, 41.0, 29.8; ESI-HRMS calcd for C₈H₁₂NO₃ [M+H] 170.0812, found 170.0812.

Figure 4.83 Preparation of aminoethylbenzene triol hydrobromide derivatives from trimethoxybenzene.
Figure 4.84 Chemical structure of trimethoxy nitrovinyl benzene (S26).

*(E)-1,2,4-Trimethoxy-3-(2-nitrovinyl)benzene (S26).* A mixture of 2,3,6-trimethoxybenzaldehyde\(^{15}\) (500 mg, 2.55 mmol, 1.00 equiv) and ammonium acetate (295 mg, 3.82 mmol, 1.50 equiv) in nitromethane (23 mL) was heated to reflux for 18 h after which the reaction was found to be according to TLC (\(R_f = 0.30\) starting material; \(R_f = 0.38\) product; hexanes:EtOAc 2:1). The resulting mixture was concentrated under reduced pressure and purified by flash column chromatography (hexanes:EtOAc 5:1) to afford nitrovinyl benzene S26 (500 mg, 82%) as a yellow solid. IR (thin film) \(\nu\) 2939, 2854, 1625, 1583, 1507, 1496, 1330, 1284, 1116, 1009 cm\(^{-1}\); \(^1\)H NMR (500 MHz, CD\(_3\)OD): \(\delta\) 8.39 (d, \(J = 13.7\) Hz, 1H), 8.09 (d, \(J = 13.7\) Hz, 1H), 7.17 (d, \(J = 9.2\) Hz, 1H), 6.80 (d, \(J = 9.2\) Hz, 1H), 3.91 (s, 3H), 3.91 (s, 3H), 3.85 (s, 3H); \(^{13}\)C NMR (125 MHz, CD\(_3\)OD): \(\delta\) 155.5, 151.6, 148.2, 140.8, 130.7, 118.7, 114.7, 107.1, 61.6, 57.0, 56.6; \(R_f = 0.38\) (hexanes:EtOAc 2:1); ESI-HRMS calcd for C\(_{11}\)H\(_{14}\)NO\(_5\) [M+H] 240.0866, found 240.0860.

Figure 4.85 Chemical structure of trimethoxyphenethyl amine (S27).

2-(2,3,6-Trimethoxyphenyl)ethan-1-amine (S27). LiAlH\(_4\) (2.0 M in THF; 3.66 mL, 7.32 mmol, 3.50 equiv) was added dropwise over 10 min to a solution of nitrovinyl benzene S26 (200 mg, 2.09 mmol, 1.00 equiv) in THF (12 mL) at 0 °C. The resulting mixture was allowed to warm to rt and stirring was continued for 24 h. The reaction mixture was cooled to 0 °C and 10% aqueous NaOH (10.0 mL) was added dropwise over 10 min, resulting in an exothermic reaction. Stirring
was stirred continued for 1 h. The resulting suspension was diluted with EtOAc (20 mL) and filtered over a plug of Celite® (EtOAc rinse). The filtrate was dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure to afford analytically pure amine S27 (438 mg, quant.) as a pale yellow oil. IR (thin film) v 3363, 2936, 2833, 1648, 1485, 1463, 1253, 1085, 793, 627 cm⁻¹; ¹H NMR (500 MHz, CDCl₃): δ 6.73 (d, J = 8.8 Hz, 1H), 6.55 (d, J = 8.8 Hz, 1H), 3.82 (s, 3H), 3.82 (s, 3H), 3.76 (s, 3H), 2.88 (d, J = 6.0 Hz, 2H), 2.83 (d, J = 6.0 Hz, 2H), 2.47 (br s, NH₂); ¹³C NMR (125 MHz, CDCl₃): δ 152.5, 148.4, 147.2, 122.7, 110.2, 105.4, 60.8, 56.2, 55.9, 42.2, 28.5; Rf = 0.10 (EtOAc:MeOH 85:15); ESI-HRMS calcd for C₁₁H₁₈NO₃ [M+H] 212.1281, found 212.1273.

Figure 4.86 Chemical structure of trimethoxyphenethyl amine (S27).

3-(2-Aminoethyl)benzene-1,2,4-triol hydrobromide (S28). Following the general phenol ether cleavage protocol described for the preparation of amine hydrobromide S25 using phenol ether S27 as starting material, triol amine S28 was obtained, according to MS identification, in low quantities along with brominated species. Initial attempts in changing the reaction temperature or the number of equivalents of BBr₃ resulted in low conversion and the desired product could not be isolated in pure form. Therefore it was not used in any enzyme assays.
4.5.3 Preparation of 2-amino-4-(2-aminoethyl)phenol

All reactions were carried out with degassed solvents under a positive pressure of nitrogen.

Figure 4.87 2-Amino-4-(2-aminoethyl)phenol via Mitsunobu strategy.

Figure 4.88 Chemical structure of benzyloxy nitrophenethyl alcohol (S29).

2-(4-(Benzyloxy)-3-nitrophenyl)ethan-1-ol (S29). (Bromomethyl)benzene (3.02 mL, 4.34 g, 25.4 mmol, 2.50 equiv) was added dropwise to a suspension of commercially available 2-(4-hydroxy-3-nitrophenyl)acetic acid (2.00 g, 10.1 mmol, 1.00 equiv), anhydrous potassium carbonate (4.21 g, 30.4 mmol, 3.00 equiv), and anhydrous potassium iodide (674 mg, 4.06 mmol, 0.400 equiv) in acetone (34 mL) at rt. Vigorously stirring was continued for 48 h. The resulting suspension was diluted with a mixture of EtOAc-H₂O (1:1; 100 mL), cooled to 0 °C, and adjusted to pH = 1 using aqueous 1M HCl. This resulted in an exothermic reaction. The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 15 mL). The combined organic
layers were washed with H₂O (2 x 15 mL), brine (30 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure to afford crude 2-(4-(benzyloxy)-3-nitrophenyl)acetic acid, which was immediately used in the next step without further purification.

BH₃ • SMe₂ (2.0 M in THF; 6.57 mL, 13.1 mmol, 1.30 equiv) was added dropwise to a solution of the crude 2-(4-(benzyloxy)-3-nitrophenyl)acetic acid in THF (110 mL) at 0 °C. The resulting mixture was allowed to warm to rt over 3 h and stirring was continued for 14 h. The resulting suspension was cooled to 0 °C and carefully quenched with the dropwise addition of saturated aqueous NaHCO₃. The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with H₂O (2 x 20 mL), brine (2 x 20 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure to afford crude alcohol S29 as a brown oil. Purification by flash column chromatography (hexanes:EtOAc 1:1) afforded analytically pure alcohol (1.98 g, 72%) as a pale yellow oil. ¹H NMR (500 MHz, CDCl₃): δ 7.71 (d, J = 2.3 Hz, 1H), 7.46 – 7.29 (m, 6H), 7.05 (d, J = 8.6 Hz, 1H), 5.19 (s, 2H), 3.82 (t, J = 6.5 Hz, 2H), 2.81 (t, J = 6.5 Hz, 2H), 1.90 (br s, OH); ¹³C NMR (125 MHz, CDCl₃): δ 150.5, 140.0, 135.8, 134.8, 131.8, 128.7, 128.2, 127.0, 125.9, 115.4, 71.3, 63.0, 37.7; Rf = 0.18 (hexanes:EtOAc 1:1). The spectral characteristics were identical to those reported in the current literature (125).

Figure 4.89 Chemical structure of benzyloxy nitrophenethyl phthalimide (S30).
2-(4-(Benzyloxy)-3-nitrophenethyl)isoindoline-1,3-dione (S30). Diethyl azodicarboxylate (DEAD; 40% in toluene; 3.95 mL, 8.05 mmol, 1.10 equiv) was added dropwise over 10 min to a solution of PPh₃ (2.21 g, 8.42 mmol, 1.15 equiv), phthalimide (1.24 g, 8.42 mmol, 1.15 equiv) and the alcohol S29 (1.98 g, 7.32 mmol, 1.00 equiv) in THF (50 mL) at 0 °C. The resulting mixture was allowed to warm to rt over 3 h and stirring was continued for 14 h. The resulting pale-yellow solution was concentrated under reduced pressure and purified by flash column chromatography (hexanes:EtOAc 4:1) to afford the desired phthalimide protected amine S30 (2.28 g, 78%) as a colorless solid. IR (thin film) ν 2985, 2871, 2783, 1774, 1750, 1640, 1387, 1307, 717 cm⁻¹; ¹H NMR (500 MHz, CDCl₃): δ 7.94 – 7.80 (m, 2H), 7.78 – 7.67 (m, 3H), 7.50 – 7.30 (m, 6H), 7.05 (d, J = 8.6 Hz, 1H), 5.20 (s, 2H), 3.91 (t, J = 7.5 Hz, 2H), 2.99 (t, J = 7.5 Hz, 2H); ¹³C NMR (125 MHz, CDCl₃): δ 168.2, 150.8, 140.1, 135.7, 134.5, 134.2, 132.0, 130.9, 128.8, 128.3, 127.1, 126.0, 123.5, 115.6, 71.3, 38.8, 33.4; Rf = 0.18 (hexanes:EtOAc 1:1); ESI-HRMS calcd for C₂₃H₁₉N₂O₅ [M+H] 403.1288, found 403.1266 and C₂₂H₁₉N₂O₅Na [M+Na] 425.1113, found 425.1111.

![Chemical structure of amino alcohol (S31)](image)

**Figure 4.90 Chemical structure of amino alcohol (S31).**

2-(3-Amino-4-hydroxyphenethyl)isoindoline-1,3-dione (S31). A flame dried round-bottomed flask was charged with nitroarene S30 (100 mg, 0.250 mmol, 1.00 equiv) in a mixture of EtOH-CH₂Cl₂ (1:1; 18 mL) at rt. Pd-C (10% on activated charcoal; 15 mg) was added to the clear solution, which was purged with H₂ for 15 min with the H₂ inlet needle below the solvent surface. The H₂ inlet needle was raised above the solvent surface and stirring was continued for 18 h. The resulting black suspension was filtered over a short plug of Celite® (CH₂Cl₂ rinse). The
filtrate was concentrated under reduced pressure to afford a brown oil that was purified by flash column chromatography (hexanes:EtOAc 1:1) to afford the desired amino alcohol S31 (70 mg, quant.) as a yellow solid. IR (thin film) ν 3373, 2941, 2824, 1410, 1022, 760, 617 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆): δ 8.77 (br s, OH), 7.96 – 7.72 (m, 4H), 6.50 (d, J = 7.9 Hz, 1H), 6.46 (d, J = 2.0 Hz, 1H), 6.20 (dd, J = 7.9, 2.1 Hz, 1H), 4.45 (br s, NH₂), 3.69 (t, J = 7.6, 2H), 2.67 (t, J = 7.6, 2H); ¹³C NMR (125 MHz, DMSO-d₆): δ 167.7, 142.5, 136.5, 134.4, 131.6, 128.9, 123.0, 116.3, 114.6, 114.3, 39.5, 33.4; Rₛ = 0.20 (hexanes:EtOAc 1:1); ESI-HRMS calcd for C₁₆H₁₅N₂O₃ [M+H] 283.1077 found 283.1066 and C₁₆H₁₄N₂O₃Na [M+Na] 305.0902, found 305.0880.

**Figure 4.91 Chemical structure of 2-Amino-4-(2-aminoethyl)phenol (S32).**

**2-Amino-4-(2-aminoethyl)phenol (S32).** Hydrazine monohydrate (120 µL, 2.48 mmol, 10.0 equiv) was added to a suspension of the phthalimide protected amine S31 (70.0 mg, 0.248 mmol, 1.00 equiv) in EtOH (1.5 mL). The resulting solution was heated to reflux for 1.5 h while colorless solids crashed out. The resulting suspension was allowed to cool to rt before H₂O (10 mL) was added in one portion. Stirring was continued to afford a clear solution that was extracted with EtOAc (3 x 5 mL). The combined organic layers were washed with H₂O (3 x 5 mL), brine (3 x 5 mL), dried over anhydrous Na₂SO₄, filtered and concentrated under reduced pressure to yield analytically pure amine S32 (18 mg, 48%). IR (thin film) ν 3384, 2947, 2822, 1580, 1239, 1049, 1021, 837 cm⁻¹; ¹H NMR (500 MHz, CD₃OD): δ 6.66 (d, J = 8.0 Hz, 1H), 6.63 (d, J = 2.2 Hz, 1H), 6.47 (dd, J = 8.0, 2.2 Hz, 1H), 3.05 (t, J = 7.5 Hz, 2H), 2.75 (t, J = 7.5 Hz, 2H), OH- and NH-protons are not visible; ¹³C NMR (125 MHz, CD₃OD): δ 145.2, 136.4, 131.2, 120.2, 117.7, 115.7,
4.5.4 Preparation of hydroxytyrosol

4-(2-Hydroxyethyl)benzene-1,2-diol (S33). LiAlH₄ (340 mg, 8.92 mmol, 5.00 equiv) was added in small portions to a solution of commercially available 3,4-dihydroxyphenylacetic acid (300 mg, 1.78 mmol, 1.00 equiv) in THF (35 mL) at 0 °C. The suspension was allowed to warm to rt over 30 min before being heated to reflux for 18 h. The resulting mixture was cooled to 0 °C and quenched with the slow addition of aqueous 0.5 M HCl (30 mL). The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 10 mL). The combined organic layers were washed with H₂O (2 x 10 mL), brine (2 x 10 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under reduced pressure to afford an orange oil. Purification by flash column chromatography (hexanes:EtOAc 1:1) afforded triol S33 (260 mg, 95%) as a pale red oil.¹H NMR (500 MHz, CD₃OD): δ 6.74 – 6.60 (m, 2H), 6.58 – 6.48 (m, 1H), 3.67 (t, J = 7.2 Hz, 2H), 2.66 (t, J = 7.2 Hz, 2H), OH-protons are not visible; Rf = 0.20 (hexanes:EtOAc 1:1). The spectral characteristics were identical to those reported in the current literature (126).
4.5.5 Additional information

2-Aminoethylbenzenediol/-triol derivatives (S16–S20, S25) as well as 2-amino-4-(2-aminoethyl)phenol S32 are sensitive towards oxidation and turn black within hours if stored in the presence of O₂. No noticeable change in their composition is observed, according to ¹H NMR, when stored in the absence of O₂ at 3–4 °C.

Salt formation of the oxygen sensitive alkylamines (S16–S20, S25, S32) leads to more stable compounds as no decomposition was observed when stored in the presence of O₂ for several days.

Upon quenching of the phenol ether cleavage reaction with MeOH, a reaction that contains BBr₃, volatile B(OMe)₃ is formed as sole side product. This can be removed under reduced pressure to afford pure products. It is important to stir the reaction mixture for approximately 1 h upon the addition of MeOH to allow for the full conversion of BBr₃ to B(OMe)₃. The obtained residue can be dissolved in additional MeOH and again concentrated under reduced pressure to ensure the complete removal of B(OMe)₃.
4.5.6 $^1$H / $^{13}$C SPECTRA

Figure 4.93 $^1$H NMR of dimethoxyphenethyl alcohol (S1).
Figure 4.94 $^{13}$C NMR of dimethoxyphenethyl alcohol (S1).
Figure 4.95 $^1$H NMR of trimethoxyphenethyl alcohol (S4).
Figure 4.96 $^{13}$C NMR of trimethoxyphenethyl alcohol (S4).
Figure 4.97 $^1$H NMR of trimethoxyphenethyl alcohol (S5).
Figure 4.98 $^{13}$C NMR of trimethoxyphenethyl alcohol (S5).
Figure 4.99 $^1$H NMR of trimethoxyphenethyl phthalimide (S10).
Figure 4.100 $^{13}$C NMR of trimethoxyphenethyl phthalimide (S10).
Figure 4.101 $^1$H NMR of dimethoxyphenethyl amine (S11).
Figure 4.102 $^{13}$C NMR of dimethoxyphenethyl amine (S11).
Figure 4.103 $^1$H NMR of dimethoxyphenethyl amine (S12).
Figure 4.104 $^{13}$C NMR of dimethoxyphenethyl amine (S12).
Figure 4.105 $^1$H NMR of trimethoxyphenethyl amine (S14).
Figure 4.106 $^{13}$C NMR of trimethoxyphenethyl amine (S14).
Figure 4.107 $^1$H NMR of trimethoxyphenethyl amine (S15).
Figure 4.108 $^{13}$C NMR of trimethoxyphenethyl amine (S15).
Figure 4.109 $^1$H NMR of aminoethylbenzene diol hydrobromide (S16).
Figure 4.110 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S16).
Figure 4.111 $^1$H NMR of aminoethylbenzene diol hydrobromide (S17).
Figure 4.112 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S17).
Figure 4.113 $^1$H NMR of aminoethylbenzene diol hydrobromide (S18).
Figure 4.114 $^{13}$C NMR of aminoethylbenzene diol hydrobromide (S18).
Figure 4.115 $^1$H NMR of aminoethylbenzene triol hydrobromide (S19).
Figure 4.116 $^{13}$C NMR of aminoethylbenzene triol hydrobromide (S19).
Figure 4.117 $^1$H NMR of nitrovinyl benzene (S23).
Figure 4.118 $^{13}$C NMR of nitrovinyl benzene (S23).
Figure 4.119 $^1$H NMR of trimethoxyphenethyl amine (S24).
Figure 4.120 $^{13}$C NMR of trimethoxyphenethyl amine (S24).
Figure 4.121 $^1$H NMR of aminoethylbenzene triol hydrobromide (S25).
Figure 4.122 $^{13}$C NMR of aminoethylbenzene triol hydrobromide (S25).
Figure 4.123 $^1$H NMR of trimethoxy nitrovinyl benzene (S26).
Figure 4.124 $^{13}$C NMR of trimethoxy nitrovinyl benzene (S26).
Figure 4.125 $^1$H NMR of trimethoxyphenethyl amine (S27).
Figure 4.126 $^{13}$C NMR of trimethoxyphenethyl amine (S27).
Figure 4.127 $^1$H NMR of benzyloxy nitrophenethyl alcohol (S29).
Figure 4.128 $^{13}$C NMR of benzyloxy nitrophenethyl alcohol (S29).
Figure 4.129 $^1$H NMR of benzyloxy nitrophenethyl phthalimide (S30).
Figure 4.130 $^{13}$C NMR of benzyloxy nitrophenethyl phthalimide (S30).
Figure 4.131 $^1$H NMR of amino alcohol (S31).
Figure 4.132 $^{13}$C NMR of amino alcohol (S31).
Figure 4.133 $^1$H NMR of 2-amino-4-(2-aminoethyl)phenol (S32).
Figure 4.134 $^{13}$C NMR of 2-amino-4-(2-aminoethyl)phenol (S32).
4.6 References


Chapter 5. Progress towards biochemical and mechanistic characterization of catechol dehydroxylases
5.1 Introduction

In Chapter 4, we described the discovery and preliminary characterization of catechol dehydroxylases, a new class of molybdenum-dependent enzymes widespread among sequenced microbes. Herein we describe progress in mechanistic characterization and heterologous expression of this chemically and ecologically intriguing enzyme class.

Though it may receive less attention than some of its neighbors in the periodic table of elements, the transition metal molybdenum is essential for most life on Earth. Since its initial discovery in the 18th century, molybdenum has emerged as a key player in our material life, from the construction industry to the synthetic laboratory. For example, this transition metal is a frequent additive in steel where it improves corrosion resistance (the pyramid at the Louvre is made from molybdenum steel). It is also a key element in chemical catalysis, enabling sulfur removal, aldehyde oxidation, and olefin metathesis (1). While human awareness and use of molybdenum is relatively recent, the biological history of molybdenum is ancient (2, 3). Molybdenum has existed since the origin of Earth, but it was tungsten, a closely related transition metal found one row below molybdenum in the period table, that was available to early life (4-6). This was because tungsten (W) sulfides (oxidation state IV), the most prevalent form of the metal on the anaerobic planet, were significantly more soluble than their molybdenum (Mo) sulfide (oxidation state IV) counterparts. However, the evolution of oxygenic photosynthesis and the resulting oxygenation of the planet’s oceans shifted this balance. Oxygenation generated highly soluble Mo (VI) salts, which became available to biology because of their increased abundance (4, 6-8). Today, the distribution of tungsten-and molybdenum-utilizing organisms reflect these early events. Organisms that require molybdenum abound across diverse ecosystems and domains of life, while organisms using tungsten as an essential nutrient are often hyperthermophilic anaerobes living at
hydrothermal vents, where tungsten is enriched and conditions mimic those on the primordial, anaerobic earth (4, 7, 9, 10).

The biological requirement for molybdenum is generally due to its involvement in essential enzymatic processes (2). This role was first proposed by Ter Meulen (11). In a seminal 1932 study, he reported the surprising presence of molybdenum in charcoal, which inspired him to search for this metal in plants, animal tissues, and environmental samples. Molybdenum appeared in many environments – from cod’s liver to Mexican crude oil. Having uncovered molybdenum in these places, he further noted the presence of exceptionally high molybdenum concentrations in the aquatic plant Azolla, which harbored a nitrogen-fixing symbiont. He also described that a separate nitrogen-fixing microbe required molybdenum for growth. Based on these data, he concluded that molybdenum is widely distributed among biological systems and proposed that molybdenum might be particularly important in nitrogen fixation (11). Since then, our knowledge of molybdenum enzymes has grown exponentially, with multiple research groups verifying Ter Meulen’s initial ideas. Notably, molybdenum forms an essential part of the heteronuclear cofactor in a type nitrogenase, the enzyme responsible for nitrogen fixation by reducing dinitrogen to ammonia (the cofactor in this enzyme type is named FeMoco) (Figure 5.1). The assembly and function of this exceptionally complex nitrogenase cofactor, which is comprised of multiple iron atoms, inorganic sulfide, and other metals, has been the subject of extensive research and remains a fascinating topic for chemists and biologists alike (12).
Figure 5.1 The molybdenum-dependent nitrogenase harbors a heteronuclear cofactor. The structure on the top shows the FeMoco cofactor present in the molybdenum-dependent nitrogenase. The reaction on the bottom represents the chemically challenging reduction of dinitrogen catalyzed by nitrogenase enzymes.
The cofactor in nitrogenase represents the exception rather than the rule among molybdenum-dependent enzymes (Figures 5.1 and 5.2). The majority of molybdenum-enzymes harbor molybdenum in complex with an organic pterin cofactor, wherein the molybdenum is coordinated by dithiolate ligand(s). These enzymes are always mononuclear, meaning they only bind one molybdenum atom, and they are generally grouped into three main families distinguished by the structure of the pterin cofactor (Figure 5.2) (13). The first family is xanthine oxidase, which binds the molybdenum atom through a single pterin. Prominent members of this enzyme family, which is present among all domains of life, include xanthine oxidase, which plays critical roles in the catabolism of purines in humans and other species (14, 15), and the microbial 4-hydroxybenzoyl CoA reductase (4-HBCR) enzyme, which degrades aromatic compounds and represented the only reported reductive aromatic dehydroxylase prior to our discovery of Dadh (see Chapters 3 and 4) (16). Sulfite oxidase represents another family of pterin-binding molybdenum-dependent enzymes (Figure 5.2). Similar to xanthine oxidases, this enzyme family binds molybdenum through a single pterin and is present throughout all domains of life. Prominent examples include the human mitochondrial sulfite oxidase, the mutation of which has fatal consequences in early life (17), and plant assimilatory nitrate reductases (18).

The exclusively microbial DMSO reductase enzyme family is the final member of the pterin-binding molybdenum-dependent enzymes. These enzymes were likely present in the last universal common ancestor (7). Whereas sulfite oxidases and xanthine oxidases bind molybdenum through a single pterin, DMSO reductase enzymes coordinate molybdenum by two dithionates from two separate pterin molecules bound to nucleotides, giving rise to the bis-molybdopterin guanine dinucleotide, or bis-MGD, cofactor (Figures 5.2 and 5.3). Because most DMSO reductase
type enzymes harbor a bis-MGD cofactor, these enzymes are also called bis-MGD enzymes, a name that we will use from here onwards.

**Figure 5.2 Mononuclear molybdenum-dependent enzymes.**
The structure on the top shows the general pyranopterin cofactor. The three structures on the bottom show the general structures of the active sites of the three major families of mononuclear molybdenum-dependent enzymes. The thiols coordinating the molybdenum come from the pyranopterin cofactor shown above. Representative reactions for each enzyme family are also shown.
Bis-MGD enzymes are exclusively found in Prokaryotes and Archaea, where they serve critical functions in metabolism. These enzymes have a wide substrate scope, ranging from organic anions to aromatic substrates, and catalyze reactions across a remarkable 1.3 V redox potential (Figure 5.3) (13). Prominent examples include respiratory DMSO reductases and nitrate reductases (19, 20), which allow the use of alternative electron acceptors; acetylene hydratase (21), which allows marine bacteria to access the inert hydrocarbon acetylene as a carbon source; pyrogallol transhydroxylase (22), which enables growth on aromatic substrates in environmental microbes; arsenate reductase (23), which enables microbial respiration on an environmental pollutant; and the catechol dehydroxylases described in Chapters 3 and 4 of this thesis, which enable gut microbial metabolism of host-and diet-derived catechols, potentially for anaerobic respiration (Figure 5.3). This diversity of transformations reflects the incredible redox versatility of molybdenum, which can access IV, V, VI, and even III oxidation states and thus serve as a bridge between 1- and 2-electron systems in biology (24). In comparison to other metallocofactors, the molybdenum atom possesses a balanced affinity for oxygen, making oxygen transfer one of the most prominent and distinguishing reactivities of bis-MGD enzymes and molybdenum-dependent enzymes more broadly (Figure 5.3) (13, 25, 26). A majority of these reactions are net redox processes.
Figure 5.3 The bis-MGD cofactor and common reactions catalyzed by family members.
A) The full structure of the bis-MGD cofactor. In rare instances, the guanine nucleotides are replaced by cytidine. B) Common reactions catalyzed by bis-MGD enzyme family members. The biological roles of the transformations are also indicated.
Researchers have used a range of tools to characterize reactions performed by bis-MGD enzymes. The molybdenum center of these enzymes absorbs across the visible spectrum and display clear spectral changes upon reduction, enabling detection of spectroscopically distinct intermediates during the course of the reaction (27). Commonly used spectroscopic methods for studying molybdenum species include resonance Raman spectroscopy, Electron Paramagnetic Resonance (EPR) spectroscopy, and X-ray absorption spectroscopy (13). Application of these tools – along with information gained from computational, biochemical, and structural studies – has shed light on the mechanism of oxygen transfer by bis-MGD enzymes (28-31). Instead of reviewing an extensive list of proposed enzyme mechanisms, we have chosen to discuss select examples that capture different types of reactivity within the broader enzyme family. Below we first discuss the proposed mechanism for the reduction of DMSO into dimethylsulfide (DMS) by *Rhodobacter sphaeroides* DMSO reductase, a deoxygenation reaction that represents a typical reactivity for bis-MGD enzymes (13). Additionally, we discuss the mechanisms of pyrogallol transhydroxylase and acetylene hydratase, two enzymes that perform chemically intriguing redox neutral oxygen transfer reactions that deviate from the typical reactivity of the enzyme family (32). Finally, drawing upon these three examples, we propose a mechanism for catechol dehydroxylation by the catechol dehydroxylase enzymes that we have described in this thesis (see Chapters 3 and 4 for details).

*R. sphaeroides* DMSO reductase is perhaps the best-characterized of all bis-MGD enzymes, with a wealth of spectroscopic, structural, biochemical, and computational work providing insight into the potential mechanism (13). The proposed mechanism for DMSO reduction is initiated by coordination of the sulfoxide oxygen to the Mo (IV) atom in the fully reduced enzyme (Figure 5.4, intermediates 2). The substrate-bound Mo (IV) species (Figure 5.4,
intermediate 2) has been observed spectroscopically and in enzyme structures \((13)\). Oxygen transfer from DMSO to the molybdenum then follows, producing an oxygen-bound Mo (VI) species (Figure 5.4, intermediate 3) and releasing the product dimethylsulfide (DMS). The precise arrow pushing for substrate reduction at this step has not been rigorously defined, but it is thought to involve 2-electron rather than stepwise 1-electron chemistry \((13)\). In the next step, oxidized Mo (VI) (Figure 5.4, intermediate 3) undergoes a single-electron reduction to generate Mo (V) intermediate that is detectable by EPR spectroscopy (Figure 5.4, intermediate 4). In the final step, the Mo (V) species is reduced to the catalytically competent (VI) species (Figure 5.4, intermediate 1), re-initiating the catalytic cycle. Studies of DMSO reductase have provided a mechanistic template for other bis-MGD enzyme catalyzing oxygen transfer reactions, including TMAO reductase, perchlorate reductase, and arsenate reductase \((33-36)\). In these enzymes, water is generally the ultimate sink of oxygen.

**Figure 5.4 Proposed mechanism for DMSO reduction by *R. sphaeroides* DMSO reductase.** The precise details of the oxygen transfer step are unclear, but it is thought that this step proceeds by a direct two-electron reduction.
Many oxygen transfer reactions catalyzed by bis-MGD enzymes involve substrate reduction coupled with the oxidation of the catalytic molybdenum atom (13). However, the reaction performed by pyrogallol hydroxytransferase represents an intriguing deviation from the redox transformations typically catalyzed by bis-MGD enzymes. This enzyme catalyzes the hydroxylation of pyrogallol into phloroglucinol, a key oxygen transfer reaction in central carbon metabolism in the marine organism Pelobacter acidigallici (37, 38). Prior to our discovery of the catechol dehydroxylases, pyrogallol transhydroxylase was the only bis-MGD enzyme known to directly modify the aromatic ring of an aromatic substrate. Early work with O¹⁸-labeled water indicated that water surprisingly was not a source of any of the hydroxys in the product (38). Instead, the hydroxyl group derives from a 1,2,3,5-tetrahydroxybenzene (THB), which serves as a co-substrate and is regenerated during the course of the reaction (38, 39). Specifically, migration of the hydroxyl group in position 2 of 1,2,3,5-tetrahydroxybenzene to position 5 of pyrogallol forms phloroglucinol and a new molecule of THB. Thus, this transformation does not represent a net redox reaction.

While a range of mechanisms have been proposed for pyrogallol hydroxytransferase, structural work suggests a mechanism in which the co-substrate THB directly participates in the reaction (Figure 5.5) (22, 39). In the proposed mechanism, the catalytically competent enzyme harbors an oxidized Mo (VI) species (Figure 5.5, intermediate 1) instead of the Mo (IV) species typically invoked in redox enzymes such as DMSO reductase. The mechanism commences with the tautomerization of the pyrogallol substrate to yield a α-hydroxyenone intermediate (Figure 5.5, intermediate 2). This step is followed by Mo-mediated oxidation to form an orthoquinone (Figure 5.5, intermediate 3). A hydroxyl group in the co-substrate THB then undergoes nucleophilic conjugate addition to orthoquinone 3, forming a covalent ether linkage to the C2 position of...
pyrogallol (Figure 5.5, intermediate 4). Further tautomeration (going through intermediates 5 and 6 in Figure 5.5) leads to a dihydroxylated para-quinone molecule bound to Mo (IV) (Figure 5.5, intermediate 7) and releases the phloroglucinol product. Mo (IV) then reduces the quinone to yield THB and the catalytically competent Mo (VI) (Figure 5.5, intermediate 8). Although analogs of the proposed bicyclic complex are accepted by the enzyme (40), biochemical and spectroscopic evidence for the proposed mechanism remain scarce. In particular, details of the steps involving electron transfer to and from the molybdenum atom remain elusive.
Figure 5.5 Proposed mechanism of hydroxy transfer by pyrogallol hydroxytransferase. This mechanism was proposed based on the crystal structure of the enzyme and studies with substrate analogs.
Acetylene hydratase represents another unusual non-redox oxygen transfer reactivity within the bis-MGD enzyme family. As described in Chapter 4, phylogenetic analysis indicated that this enzyme is the biochemically characterized enzyme most closely related to the catechol dehydroxylases. Acetylene hydratase catalyzes the hydration of acetylene, producing acetaldehyde that can be shuttled into central carbon metabolism and support microbial growth (21). This enzyme harbors tungsten in the active site but also functions with molybdenum (21). Surprisingly, this enzyme also harbors a redox-active N-terminal 4Fe-4S cluster despite performing a redox neutral reaction (41, 42). Mutagenesis of lysine at position 48 in acetylene hydratase, a conserved residue involved in the electron transfer from the 4Fe-4S cluster to W or Mo in bis-MGD enzymes, did not affect catalysis, raising questions about the role of this cluster in this non-redox enzyme (43). This cluster may modulate the pKa of an active site aspartate residue (Asp 13) that is critical for reactivity (41).

The currently proposed mechanism for acetylene hydration is based on quantum chemical calculations, a crystal structure where the substrate was manually docked in the active site, and mutagenesis studies (Figure 5.6) (32, 42-47). Acetylene hydratase activity requires a strong reductant, suggesting W(IV) as the oxidation state during the course of the non-redox reaction (21). The crystal structure suggests that water is coordinated to the W(IV) in the catalytically competent enzyme (Figure 5.6, intermediate 1) (41). In the first step of the reaction, the acetylene substrate displaces the water molecule and coordinates to W(IV), forming intermediate 2 in Figure 5.6. The displaced water molecule is then deprotonated by an active site residue and undergoes nucleophilic addition to acetylene. This results in a W(IV)-coordinated vinyl anion (Figure 5.6, intermediate 2), which then deprotonates the active site residue to form an enol (Figure 5.6, intermediate 4). Next, the enol hydroxyl group is de-protonated, enabling in its direct coordination.
to the W (IV) metal (Figure 5.6, intermediate 5). Non-enzymatic tautomerization produces acetaldehyde, which escapes the active site and makes the W(IV) species available for coordination (Figure 5.6, intermediate 6). Water binding to W (IV) reinitiates the catalytic cycle. Mutagenesis studies have implicated an aspartate at position 13 as the catalytic base performing the key water activation and tautomerization steps in this reaction (43). Changing Asp 13 into glutamate did not change enzymatic activity, while mutagenesis into an alanine effectively abolished activity, highlighting the importance of the carboxylate group (43). However, further experimental work is required to provide supporting evidence for the proposed mechanism. For example, obtaining a substrate-bound crystal structure to determine the mode of substrate binding and the protonation states of active site residues will provide valuable mechanistic information.
Figure 5.6 Proposed mechanism of acetylene hydration by acetylene hydratase.
This mechanism was proposed based on the crystal structure of the enzyme, mutagenesis, and computational calculations. The catalytic base is likely an aspartate at position 13.
Catechol dehydroxylation potentially represents a new type of deoxygenation reaction within the bis-MGD enzyme family. Our proposed mechanism for catechol dehydroxylation invokes elements of the aforementioned examples of bis-MGD enzyme-catalyzed reactions to explain this unique reactivity (Figure 5.7). We preliminarily propose a mechanism in which the dopamine $p$-hydroxyl group coordinates to the reduced Mo(IV) center (Figure 5.7, intermediate 1). Substrate binding could be followed by tautomerization of the $m$-hydroxyl group of intermediate 2 in Figure 5.4 to a ketone with protonation of the adjacent carbon atom. In the final step, oxygen atom transfer to molybdenum could be accompanied by rearomatization of the substrate (Figure 5.7, intermediate 3) to yield the dehydroxylated product (Figure 5.7). Concurrently, Mo(IV) is oxidized to Mo(VI). In order to regenerate the active Mo(IV) species, oxygen-bound Mo(VI) (Figure 5.7, intermediate 4) would then be reduced to Mo(IV) through sequential one-electron reductions, going through the hydroxylated Mo(V) intermediate (Figure 5.7, intermediate 5) that has been observed in other bis-MGD enzymes catalyzing net redox oxygen transfer reactions [13]. Incorporation of a single proton into the product, from the solvent or surrounding amino acids, is a key feature of this mechanism. Our mechanistic proposal is consistent with the requirement of a catechol for both Dadh and Hcdh, as described in Chapter 4. Additionally, the proposed key tautomerization steps and acid-based chemistry resemble the water activation and tautomerization proposed for acetylene hydratase [42], the biochemically characterized enzyme most closely related to catechol dehydroxylases.

Catechol dehydroxylation may represent a new strategy for aromatic dehydroxylation. Our proposed mechanism and our studies of Dadh substrate scope described in Chapter 5 suggest mechanistic distinction between catechol dehydroxylases and the only other biochemically characterized reductive aromatic dehydroxylase, 4-HCBR [16]. 4-HCBR is a distinct molybdenum
dependent-enzyme belonging to the xanthine oxidase enzyme family. This enzyme contains a monomeric molybdopterin cofactor and uses a Birch reduction-like mechanism to remove a single aromatic hydroxyl group from 4-hydroxybenzoyl CoA. While 4-HCBR requires an electron-withdrawing thioester group to stabilize radical anion intermediates (16), we have found that catechol dehydroxylation does not require an electron-withdrawing substituent and can tolerate additional electron-donating hydroxyl groups on the aromatic ring (see description of Dadh substrate scope in Chapter 4). This tolerance of electron donating group, requirement for a catechol, and the observation that 4-HCBR belong to entirely different enzyme classes, supports our proposal that the enzymes use distinct mechanism for reductive aromatic dehydroxylation. It is also worth noting that catechol dehydroxylases dehydroxylases are phylogenetically distinct from pyrogallol hydroxytransferase (see Chapter 4 for details) and that our proposed mechanism also differs from that of this enzyme, which catalyzes a net redox neutral reaction. This illustrates that catechol dehydroxylation may represent a new mechanism for aromatic dehydroxylation in nature. However, further studies using spectroscopy and substrate analogs are necessary to validate our mechanistic proposal.
The proposed mechanism for catechol dehydroxylation involves de-aromatization and C–O bond cleavage driven by re-aromatization. The molybdenum atom is likely coordinated by an active site amino acid residue (here indicated as X, likely Aspartate or Cysteine), which is displaced upon catechol binding.

**Figure 5.7 Proposed mechanism for catechol dehydroxylation by catechol dehydroxylases.**
The biogenesis of molybdenum-dependent enzymes is complex and energetically expensive. This is apparent by considering the biosynthesis of bis-MGD cofactor, a process that involves at least 17 different proteins (Figure 5.8) (48). Biosynthesis begins with guanine triphosphate (GTP), which is converted to cyclic pyranopterin monophosphate (cPMP) by a radical SAM enzyme-dependent process (49, 50). The cofactor then undergoes sulfuration, followed by phosphate ester hydrolysis to create the molybdopterin core (named MPT) (51). The subsequent steps involve attachment of AMP to create MPT-AMP, and hydrolytic release of adenosine and metal incorporation by MoeA (51, 52). Importantly, MoeA can incorporate both molybdenum (as molybdate) and the closely related metal tungsten (as tungstate) at this step (52). Molybdenum (Mo) and Tungsten (W) have similar atomic radii (1.75 Å and 1.78 Å respectively), electronegativity, and covalent solution radii (4). However, redox reactions catalyzed by tungsten typically occur at much lower potentials than those catalyzed by molybdenum. As a result, some enzymes can function with either metal, while others are only functional with either Mo or W (21, 53, 54). Because MoeA does not distinguish between tungsten and molybdenum and these metals are imported by similar transporters, the type of metal that is incorporated largely depends on environmental availability (53, 55).

The formation of metal-bound MPT represents a branching point between enzyme classes; sulfite and xanthine oxidases use this cofactor without major modifications, while the microbial bis-MGD enzymes require further processing to produce the mature cofactor (13). In the final two steps of bis-MGD biosynthesis, MobA first dimerizes two MPT molecules around the central metal ion, followed by ligation of a nucleotide to form the final dinucleotide cofactor (56, 57). Guanine is the most commonly incorporated nucleotide, but cytosine can be incorporated as well (56, 58). The conservation of most steps of molybdopterin biosynthesis from microbes to humans testifies
to the critical role of molybdenum in biology (59).

Figure 5.8 Biosynthesis of the bis-MGD cofactor.
The figure shows the major steps and proteins involved in the full biosynthesis. At the step of metal incorporation, MoeA can incorporate the oxyanions tungstate (W) or molybdate (Mo).
Insertion of the full bis-MGD cofactor and subsequent maturation of the molybdenum enzyme is a complex process involving multiple interacting proteins (60). The bis-MGD cofactor is generally inserted by chaperones encoded by genes proximal to the gene encoding the bis-MGD enzyme. These chaperones are often system-specific and insert the cofactor into only their dedicated partner enzyme (61-64). Perhaps the best-understood example is the assembly of the multi-subunit *E. coli* DMSO reductase, an enzyme that comprises the bis-MGD binding catalytic subunit DmsA, an electron shuttling ferredoxin DmsB, and the membrane anchor DmsC. The dedicated cofactor insertion chaperone is named DmsD. Along with other general chaperones in the cell, DmsD stabilizes the nascent peptide chain of the catalytic subunit DmsA as it exits the ribosome (65). DmsD then interacts with enzymes catalyzing terminal steps of the bis-MGD biosynthesis pathway and facilitates cofactor insertion by an unknown mechanism. In catalytic subunits such as DmsA, which harbor an N-terminal 4Fe-4S cluster, the cluster has to be successfully assembled for successful bis-MGD insertion (20, 66). In the final steps of enzyme maturation, the folded and mature DmsA interacts with its ferredoxin partner DmsB, which harbors four fully mature 4Fe-4S clusters. The fully folded DmsAB complex is translocated across the membrane by the twin-arginine translocation (TAT) system (67, 68). The TAT signal is only present on the N-terminus of the catalytic subunit, suggesting that the ferredoxin “hitchhikes” during transport (68, 69). As described in Chapter 4, the *E. lenta* dehydroxylases harbor a TAT signal in the catalytic subunit.

The TAT signal on DmsA is cleaved following the membrane translocation event (69). The DmsAB complex then associates with the membrane anchor DmsC, a process that requires the intact C-terminus of DmsB (70). Notably, proper bis-MGD insertion is critical for the proper localization of the holo enzyme in the cell (71), highlighting that the importance of metallocofactor
installation extends beyond enabling catalytic activity. The complex assembly of bis-MGD enzymes and the requirement for multiple interacting protein complexes, some of which may only be present in the native organism, presents large challenges for heterologous expression of these enzymes.

Bis-MGD enzymes have varying architectures and subunit compositions (Figure 5.9) (24). The archetypal bis-MGD enzymes harbor three different subunits: a membrane anchor, which can either bind heme or not harbor any cofactors, an electron shuttling ferredoxin subunit, which harbors four distinct 4Fe-4S clusters, and the catalytic subunit, which usually contains an N-terminal 4Fe-4S cluster (named the F0 cluster) in addition to the bis-MGD cofactor. In the canonical respiratory nitrate and DMSO reductases from *E. coli*, electrons travel from quinones to the membrane anchor through the 4Fe-4S clusters of the ferredoxin, ultimately reducing the F0 cluster and then the molybdenum atom in the catalytic subunit (24, 72). The electrons are then transferred from the molybdenum atom to the substrate. This process is usually coupled to proton translocation across the membrane (73), providing an electron sink during respiration. As described in Chapter 4, the newly discovered dehydroxylases from *E. lenta* likely possess the archetypal three-subunit composition.

Other bis-MGD enzymes harbor unique combinations of the three subunits. For example, acetylene hydratase harbors only the catalytic subunit with the bis-MGD cofactor and the N-terminal F0 cluster (43), while the *E. coli* trimethylamine-N-oxide (TMAO) reductase (35) and the *Gordonibacter* dehydroxylases described in Chapter 4 possess a catalytic subunit without an F0 cluster. Other notable deviations from the archetypal composition include pyrogallol transhydroxylase (22), which harbors only the bis-MGD cofactor in the catalytic subunit and binds three 4Fe-4S clusters in the ferredoxin subunit. Arsenite oxidase also deviates from the archetypal
bis-MGD enzyme subunit composition. (74). The catalytic subunit of this enzyme harbors the bis-MGD cofactor and an N-terminal 3Fe-4S cluster, but the associated ferredoxin has only a single cluster of the non-typical 2Fe-2S Rieske type. Importantly, the cellular location also varies between different enzymes depending on the presence of a signal sequence (Figure 5.9). The modular composition of bis-MGD enzymes allows for careful tuning of redox potentials and chemical reactivity to serve unique physiological purposes.
Figure 5.9 Subunit composition and cellular location of microbial bis-MGD enzymes. The subunits are not drawn to scale.
In addition to the subunit composition, characteristics of the enzyme active site influence reactivity in bis-MGD enzymes. This includes the ligands that directly coordinate the catalytic molybdenum atom (Figure 5.2). In the fully mature *holo* enzyme, the central metal atom is coordinated by a terminal non-amino acid ligand (sulfido, seleno, or oxido) and an amino acid residue (usually serine, aspartate, cysteine, or seleocysteine) (Figure 5.2) (13). Early studies of the *E. coli* formate dehydrogenase Fdh suggested an important role for the amino acid ligand in modulating reactivity (75). In Fdh, the molybdenum atom is coordinated by a selenocysteine. Mutation of this residue to a cysteine dramatically decreased the $k_{cat}$ of formate oxidation by several orders of magnitude without substantially changing the $K_m$. This dramatic loss in activity could either result from indirect changes to the molybdenum coordination sphere or from direct participation of the selenocysteine in catalysis. Further investigation supported the latter hypothesis, revealing that mutation to cysteine changed the pH dependence of activity rather than abolishing activity altogether. Specifically, the lower pKa of selenocysteine ($pK_a = 5.2$) compared to cysteine ($pK_a = 8.3$) makes this residue more suitable for a critical proton abstraction at physiological pH (76). While this example highlights the importance of the amino acid ligand in enzyme activity, other studies have revealed no clear relationship between amino acid ligand identity and reactivity across bis-MGD enzymes. For example, the two evolutionarily distinct nitrate reductases NarG and NapA/NarB have different amino acid ligands despite performing the exact same chemical reaction (24). This suggests that the identity of the amino acid ligand likely matters most if this residue directly participates in catalysis.

In contrast to the amino acid ligand, the terminal non-amino acid ligand influences catalysis by modulating the coordination sphere. This ligand is most commonly sulfido or oxido. The oxido ligand is generated by default as part of cofactor biosynthesis (Figure 5.8), but the sulfido ligand
is generated by dedicated sulfuration chaperones (77, 78). Ligand identity cannot be predicted bioinformatically, and differences in enzyme preparations used for crystal structures and spectroscopy has complicated efforts to determine ligand identity in the native, active enzyme (13, 35, 79-84). For example, the crystal structure of the *E. coli* TMAO reductase indicate presence of an oxido ligand (80). Yet, TMAO reduction is highly sensitive to oxygen (85). Because this enzyme does not harbor any oxygen-labile iron sulfur clusters, this observation has suggested the presence of a sulfido ligand, which oxygen exposure could non-enzymatically convert to a less active oxo ligand.

To determine the true identity of the terminal ligand in TMAO reductase, Leimkuhler and co-workers recently developed an in vitro system for biosynthesis of the full bis-MGD cofactor and its chaperone-assisted incorporation into the *apo* TMAO reductase (35). Using this system, they revealed that incorporation of a bis-MGD cofactor harboring a sulfido ligand yielded a fully active protein whereas incorporation of an oxido ligand cofactor produced inactive enzyme. Chemical re-sulfuration of the oxido ligand using sodium sulfide restored enzymatic activity, while destroying the sulfido ligand with cyanide or oxygen ablated activity, confirming a causal role for this ligand in TMAO reduction. Cyclic voltammetry suggested these effects were directly due to the modulation of the molybdenum coordination sphere, as the redox potential of the oxo ligand protein was more positive than the sulfido ligand counterpart (35). Based on these results, the authors proposed a new mechanism of TMAO reduction and suggested that the non-amino ligand directly changes the character of the redox active metal $d$-orbitals in bis-MGD enzymes more generally (35). In addition to revealing a unique mode of tuning chemical reactivity, these results raise questions about identity of the terminal non-amino acid ligand in many bis-MGD enzymes.
that, like TMAO reductase, have been crystallized aerobically. Exposure of these enzymes to oxygen would inadvertently generate an oxo ligand.

The three-dimensional conformation of the metal-ligand complex also influences enzyme reactivity. Many molybdenum-containing organometallic complexes have been constructed to mimic the reactivity of bis-MGD enzymes \((26, 86)\). While these organometallic complexes typically exhibit low turnover rates and poor catalytic efficiency compared to their enzyme counterparts, structural comparison of these synthetic molecules with enzyme active sites have provided new insight into the factors modulating enzyme activity. For instance, a synthetic mimic of the DMSO reductase active site has a square-pyramidal geometry at the molybdenum center, while the ligand-metal complex in DMSO reductase adopts a surprising trigonal prismatic conformation that is not predicted by valence shell electron pair repulsion (VESPR) theory \((29, 86)\). Computational density functional theory (DFT) calculations modelling the DMSO reductase active site suggested that the non-VESPR trigonal prismatic conformation carries functional significance \((87)\). Key transition states during DMSO reduction also harbor a trigonal prismatic geometry, indicating that the unique pterin conformation in DMSO reductase pre-organizes the active site to lower the energy barrier for catalysis \((87)\). This prediction is consistent with spectroscopic studies of DMSO reductase, which has observed the intermediates proposed by the DFT calculations \((13, 29, 88)\). Thus, tuning the three-dimensional conformation of the ligand-metal complex represents an additional strategy by which bis-MGD enzymes modulate reactivity.

The chemical structure of the pterin backbone also influences cofactor reactivity \((89, 90)\). The two pterins that coordinate the catalytic metal in bis-MGD enzymes are differentially positioned relative to the F0 Fe-S cluster. In the majority of bis-MGD enzymes, the distal pterin \((Q)\) is structurally identical to the proximal pterin \((P)\) closest to the F0 cluster. However, in some
enzymes, the normally tricyclic Q pterin has opened, creating a bicyclic pterin (89). This ring-opened structure was originally considered a crystallographic artefact when it was first observed in the nitrate reductase NarG structure (91), but discovery of the same bicyclic structure in the phylogenetically related enzymes perchlorate reductase (Pcr) and ethylbenzene dehydrogenase (Ebd) hinted at a potential functional significance (34, 92). Further analysis of the NarG crystal structure confirmed this suspicion, revealing three amino acids (one serine and two histidines) that are poised to stabilize the open conformation of the Q pterin (93). Mutagenesis of these residues did not drastically change the EPR spectrum of the Mo(V) intermediate, indicating no major effect on the metal coordination sphere. However, changing the two histidine residues to alanines substantially decreased the overall redox potential of the enzyme, as assessed by cyclic voltammetry (93). This suggested that the two histidine residues form part of an electron transfer relay. By connecting the pyran ring oxygen to a deeply buried water molecule, this electron transfer relay system directly modulates the molybdenum redox potential and influences reactivity (93).

Pterin structure and molybdenum ligand identity alone cannot account for the vast diversity of reactions catalyzed by bis-MGD enzymes. Because bis-MGD cofactor is deeply buried in the protein, substrate must generally enter through deep substrate funnel. The amino acid composition of the funnel represents an additional means of modulating reactivity and selectivity. For example, in acetylene hydratase, mutation of an isoleucine at position 142, a residue that forms part of a hydrophobic pocket at the end of the active funnel but is not expected to directly participate in catalysis, results in a dramatic loss of activity. Similarly, recent comparisons of nitrate reductase (NarG) and perchlorate reductase (PcrA), two closely related enzymes that share similar overall folds as well as the same ligand identity and open bicyclic pterin structure (34), have reinforced the importance of the active site tunnel in reactivity. Both enzymes reduce small inorganic anions
(nitrate and perchlorate). However, purified PcrA turns over perchlorate at an order of magnitude higher catalytic efficiency compared to NarG, whereas purified NarG turns over nitrate at roughly ten-fold higher efficiency than PcrA (34). The PcrA crystal structure unveiled some notable exceptions to the highly conserved residues in the substrate funnels, including residue 461, which is a tryptophan in PcrA but is a glutamate in NarG (34). Mutation of the PcrA tryptophan into glutamate altered the catalytic efficiency towards perchlorate to closely resemble that of NarG (34), indicating an important role of this aromatic residue in gating access to the active site and regulating substrate specificity.

Structural comparisons have further highlighted the importance of active site access. For instance, NarG has a relatively narrow substrate funnel and is only active towards small oxyanions, whereas the \textit{Rhodobacter} DMSO reductase has a wide substrate funnel and turns over a large diversity of sulfoxides beyond its native substrate DMSO (94, 95). Notably, acetylene hydratase, the biochemically characterized enzyme most closely related to catechol dehydroxylases, has a unique substrate funnel where substrate enters from the N-terminus, raising the possibility that the related dehydroxylases might harbor a similarly unique gating mechanism (41). However, because the location of the substrate funnel cannot be predicted by bioinformatics, further structural and biochemical work is necessary to understand the precise amino acids that determine the active site access and substrate scope of bis-MGD enzymes.

Structurally and biochemically characterizing bis-MGD enzymes is not a trivial task because of the challenges associated with obtaining the high amounts of proteins needed for commonly used assays (34, 35). Additionally, the complex process of bis-MGD enzyme maturation and the intricacies of the cofactor complicates the use of in vitro reconstitution and heterologous expression techniques typically used to improve the yield and activity of other enzymes.
metalloenzyme families (96, 97). A majority of well-characterized bis-MGD enzymes, including but not limited to acetylene hydratase, perchlorate reductase, pyrogallol transhydroxylase, have been accessed via purification from the native organism. In these cases, organisms were cultured in minimal media where enzyme expression enabled healthy growth by providing access to electron acceptors or carbon substrates (21, 22, 34). Such conditions increase the relative expression of the enzyme of interest in comparison to other proteins, enabling purification of proteins in sufficiently high yield and purity for crystallographic, spectroscopic, and biochemical studies.

Native purification represents one of the oldest methods for protein purification and its utility remains high even with currently available molecular biology tools. Nonetheless, this purification method carries some important limitations. First, the physiological purpose of a specific transformation is not always clear, which makes it challenging to identify culture conditions that produce protein in sufficient quantities. Additionally, mutagenesis and the addition of tags for affinity chromatography are not possible when proteins are natively purified from genetically intractable organisms, as in the case of the Gordonibacter and Eggerthella dehydroxylases. Alternatively, in genetically tractable organisms, the protein can be overexpressed on a plasmid under the control of a strong promoter. This strategy was successfully employed in the purification of E. coli nitrate and DMSO reductases, as well as in the recent biochemical and structural characterization of the respiratory arsenate reductase from Shewanella sp. ANA-3 (20, 33, 34). Development of genetic tools and expression systems in organisms harboring high numbers of bis-MGD enzymes, such as the catechol dehydroxylating gut microbes Eggerthella and Gordonibacter (89, 98), would significantly accelerate enzyme discovery and characterization.
Only a few successful examples of heterologous expressed bis-MGD enzymes have been reported. In 2008, Newman and co-workers reported purification of the *Shewanella* sp. ANA-3 respiratory arsenate reductase (Arr) from an *E. coli* expression system (99). The authors first attempted expression of the catalytic subunit ArrA on a pET32H vector derivative using a C-terminal His$_6$ tag. However, this resulted in poor binding to the nickel column. Thus, the researchers replaced the first 41 residues comprising the ArrA TAT signal sequence (which was predicted to be cleaved off during enzyme maturation) with a His$_9$ tag. ArrB, the predicted ferredoxin partner, was also added to the construct C-terminally to the catalytic ArrA subunit. The final N-His$_9$ ArrAB construct was expressed in *E. coli* C43 grown in rich LB medium supplemented with 20 mM lactate and 1 mM molybdate. Ultimately, nickel affinity chromatography and size exclusion chromatography from 8 liters of culture yielded low levels of a 131 kDa ArrA-ArrB complex. No activity was observed when ArrA and ArrB were expressed individually, supporting the idea that complex formation is necessary for proper enzyme maturation and biochemical activity. These efforts not only provided the first biochemical evidence that Arr catalyzes arsenate reduction but also inspired the high-yielding expression of ArrAB in the native producer *Shewanella* sp. ANA-3 on a broad host-range pBAD18K vector, which enabled further structural, spectroscopic, and biochemical studies (33).

Acetylene hydratase, the biochemically characterized enzyme most closely related to the catechol dehydroxylases described in Chapter 4, was also heterologously expressed in *E. coli* (43). This enzyme was initially natively purified from the genetically intractable metabolizing organism *P. acetylenicus* to enable structural characterization (21, 41). To facilitate mutagenesis, the gene was cloned into a pET24 vector with a C-terminal His$_6$ tag and expressed in *E. coli* Rosetta, a strain equipped with many rare amino acid codons. Expression in a minimal mineral medium
containing 0.5 % glycerol as the carbon source and 50 mM fumarate as the electron acceptor enabled purification of a functional acetylene hydratase. However, the heterologously expressed enzyme had significantly lower levels of tungsten and iron than the natively purified acetylene hydratase, prompting the researchers to reconsider their cloning strategy. To improve metal incorporation, the authors added the first 108 nucleotides of the respiratory *E. coli* nitrate reductase NarG to the N-terminus of acetylene hydratase. These nucleotides encode for the TAT signal sequence that is critical for proper cofactor insertion and protein maturation of NarG, providing a potential means of guiding endogenous *E. coli* machinery to interact with the acetylene hydratase. This construct produced protein with tungsten levels that were significantly higher than without the added NarG TAT sequence, approaching the levels observed in the natively purified protein. The NarG-acetylene hydratase also harbored 86% of the iron of the natively purified protein, highlighting that co-opting endogenous bis-MGD enzyme maturation machinery can be critical for successful heterologous expression (43).

In Chapters 3 and 4 of this thesis we have described the discovery of the catechol dehydroxylases. While our assays with natively purified Dadh and Hcdh has provided insight into the substrate scope of gut bacterial dehydroxylases, many questions remain unanswered. For instance, even though we have proposed a mechanism for catechol dehydroxylation based on the enzyme substrate scope and known mechanisms of other bis-MGD enzyme-catalyzed reactions (Figure 5.7), further experimental work is needed to evaluate our hypothesis. Additionally, the molecular basis of the striking dehydroxylase specify is unclear. For example, how could Dadh described in Chapter 4 distinguish between dopamine and the closely related substrate hydroxytyrosol, in which the amine has been replaced by a hydroxyl group? Our current enzyme preparation is limited in purity and yield, preventing the use of EPR spectroscopy, kinetic assays,
protein crystallography, and other commonly used approaches to answer these outstanding questions about catechol dehydroxylase reactivity. This limitation highlights a need for heterologous expression systems, which could significantly improve the protein purification and enable critical mutagenesis studies. Such systems could also help de-orphan the functions of the diversity of uncharacterized enzymes within the catechol dehydroxylase enzyme class. Below we describe preliminary experiments to probe the mechanism of catechol dehydroxylation as well as efforts to heterologously express a range of catechol dehydroxylases across different conditions. At the end of this chapter, we provide experimental suggestions for further heterologous expression efforts.

5.2 Results

5.2.1 Lysate assays reveal that hydrocaffeic acid dehydroxylation incorporates high levels of a single deuterium into the product m-hppa

As described above, we have proposed a preliminary mechanism for catechol dehydroxylation by the newly discovered catechol dehydroxylases (Figure 5.7). To investigate the plausibility of this proposed mechanism, we explored the dehydroxylation of hydrocaffeic acid in deuterated water (D$_2$O). We used cell lysates instead of purified enzymes as proxies for biochemical activity because of the laborious processes associated with obtaining the pure Hcdh protein. Based on the predicted mechanism, which involves an incorporation of a single hydrogen into the product (Figure 5.7), we predicted that hydrocaffeic acid dehydroxylation, when performed in D$_2$O, would yield m-hppa with a single deuterium incorporated into the aromatic ring (Figure 5.10).

To evaluate this hypothesis, we grew *E. lenta* A2 with and without hydrocaffeic acid and then washed the cells twice in either buffered water or D$_2$O (50 mM Tris pH 8), followed by lysis
and a 4-hour incubation period to allow for potential deuterium exchange. We then exposed the lysates to hydrocaffeic acid and analyzed the reaction products by LC-MS. Strikingly, the experiment run in D₂O yielded the dehydroxylated product with >95% incorporation of a single deuterium, with the +1 mass shift being absent from reactions performed in non-deuterated buffer or in deuterated buffer exposed to oxygen (which inactivate the enzyme) (Figure 5.10). High-resolution MS/MS analysis suggested that the deuterium was incorporated at the aromatic ring, consistent with our mechanistic proposal (Figure 5.11). Though the exact position awaits further verification by Nuclear Magnetic Resonance (NMR) spectroscopy, we speculate that the deuterium replaces the hydroxyl group during the reaction (Figure 5.11). The high deuterium incorporation hints at a potential involvement of a water molecule in the protonation of the aromatic ring – an intriguing possibility given that the reaction catalyzed by the related acetylene hydratase involves a water molecule. However, further studies using spectroscopy and substrate analogs are necessary to validate our mechanistic proposal. Such mechanistic work will require further optimization of the protein preparation, which currently is limited in its yield and overall purity. Below we describe progress towards this goal, outlining our efforts to heterologously express catechol dehydroxylases from gut and environmental bacteria.
Figure 5.10 Lysate assays reveal that hydrocaffeic acid dehydroxylation incorporates high levels of a single deuterium into the product m-hppa

A) Predicted incorporation of hydrogen and deuterium into m-hppa based on proposed mechanism for hydrocaffeic acid dehydroxylation. B) Schematic describing the experimental setup to probe deuterium incorporation. E. lenta A2 was grown in BHI medium containing 1% Arginine and 10 mM formate with or without 0.5 mM hydrocaffeic acid. Cells were harvested after approximately 16 hours and were resuspended and washed three times in 50 mM Tris pH 8 made in either deuterated or non-deuterated water. Cells were lysed and were exposed to substrate following an incubation period to allow deuterium exchange. Reaction mixtures contained 500 µM catechol substrate, 500 µM methyl viologen, and 1 mM sodium dithionite, and were incubated at room temperature for 24 hours. Reactions were performed under anaerobic conditions unless otherwise indicated. C) Results from lysate assays. The reactions run in deuterated water incorporated >95% deuterium. The enzyme reactions were analyzed by LC-MS/MS. Bars represent the mean ± the SEM of three independent enzyme reactions.
Figure 5.11 High-resolution LC-MS/MS indicates that a single deuterium is specifically incorporated into the aromatic ring of m-hppa.

The reactions producing either deuterated or non-deuterated product in Figure 4.32 were analyzed by high-resolution LC-MS/MS to narrow the location of deuterium incorporation. A) MS/MS fragmentation pattern for the non-deuterated product. B) MS/MS fragmentation pattern for the deuterated product.
5.2.2 General approach for heterologous expression

In all expression experiments described below we used SDS-page analysis to assess expression and used LC-MS to analyze enzymatic activity in anaerobic cultures and lysates. Bis-MGD enzymes cannot be non-enzymatically reconstituted with the bis-MGD cofactor in vitro, meaning that if there is no enzymatic activity in lysates or cultures, there is no chance to non-enzymatically in vitro reconstitute this activity with a purified protein. Consequently, gain of enzymatic activity in lysates or cultures of the heterologous host served as a readout of proper cofactor installation and/or enzyme maturation. Additionally, all co-transformed plasmids used in our studies carried compatible origins of replication.

5.2.3 Expression of Dadh in E. coli

We initially focused on heterologous expression of the molybdenum-dependent dopamine dehydroxylase Dadh in E. coli. We first analyzed the genomic context surrounding dadh, focusing on the genes that were upregulated in response to dopamine exposure (described in Chapter 3) (Figure 5.12). The catalytic subunit, which we call dadhA, co-localizes with its predicted ferredoxin (dadhB) and membrane anchor (dadhC) partners. In addition to the three-gene operon encoding for DadhABC, we identified an additional, separate operon that harbors three ferredoxins (dadhFGH), a putative system-specific bis-MGD insertion chaperone (dadhE), and a protein of unknown function (dadhD). We further analyzed DadhA, identifying the cysteine residues predicted to coordinate its conserved N-terminal F0 4Fe-4S cluster. We also identified an N-terminal TAT signal sequence that targets this protein for translocation across the cell membrane and likely serves as a recognition site for bis-MGD-inserting chaperones (Figure 5.12). The TAT signal sequence would be cleaved off as part of normal protein processing, as we observed with the natively purified Hcdh described in Chapter 4.
Figure 5.12 Analysis of *dadh* and its surrounding genomic context.
A) Analysis of the genomic context of Dadh reveals the existence of two operons, one encoding the DadhABC subunits that comprise the Dadh enzyme, and the other harboring other genes, including ferredoxins and the predicted bis-MGD insertion chaperone DadhE. B) The first 34 amino acids of DadhA are predicted to encode a TAT signal sequence, which is likely cleaved between positions 34 (alanine) and 35 (glutamate) as part of normal enzyme maturation and processing. C) DadhA harbors conserved cysteine residues (at positions 9, 12, 16, and 46) predicted to coordinate the N-terminal F0 cluster. DadhA has been aligned with acetylene hydratase (Ach), a structurally and biochemically characterized enzyme that also harbors an F0 cluster. The alignment has been truncated to highlight the positions of the cysteines.
Following this analysis, we cloned \textit{dadhA} variants differing in the location of the His\textsubscript{6} tag (C-terminal or N-terminal) and whether the \textit{dadhA} TAT signal sequence was truncated (Table 5.1). We used the pET28a plasmid (driven by the strong T7 promoter) in the commonly employed \textit{E. coli} BL21 host strain (DE3). This system had previously yielded high quantity and purity of active TyrDC and AADC, as described in Chapter 2. We supplemented the LB expression medium with nitrate, molybdate, and iron to facilitate anaerobic respiration and bis-MGD enzyme production. Comparing aerobic and anaerobic expressions revealed that protein levels varied between constructs, an observation that was likely explained by interferences with \textit{dadhA} TAT signal sequence recognition by \textit{E. coli} (Figure 5.13 and Table 5.1). For example, whereas full-length C-terminally His-tagged DadhA (Dadh-1) was clearly expressed, the C-terminally His-tagged Dadh lacking the endogenous TAT signal sequence (Dadh-3) did not appear to be expressed (Figure 5.6 and Table 5.1). Similarly, the N-terminally tagged full-length DadhA, where the tag was added in front of the endogenous TAT sequence, yielded only high levels of insoluble protein (Dadh-4), potentially because the N-terminal His\textsubscript{6} tag interfered with assembly and processing (Figure 5.13 and Table 5.1). In addition to assessing expression, we measured activity by adding the substrate dopamine to anaerobic cultures and lysates. No \textit{m}-tyramine production was observed using LC-MS, suggesting all constructs yielded inactive protein.

Our initial expression indicated that the TAT signal sequence was important for soluble expression. Even though this signal sequence could plausibly be recognized by the highly conserved TAT machinery in \textit{E. coli}, the lack of interacting chaperones could lead to improper bis-MGD insertion and explain the lack of activity. To overcome this limitation, we cloned two additional constructs. In one case, we replaced the endogenous DadhA TAT signal sequence with the TAT sequence from the \textit{E. coli} respiratory nitrate reductase NarG (creating Dadh-5), a strategy
that had previously yielded pure and active heterologously expressed acetylene hydratase (43) (Table 5.1). In another case, we cloned the predicted chaperone dadhE, into a pETDuet1 vector with a N-terminal His₆-tag (construct 7.1 in Table 5.1), co-expressing this construct with the full-length and truncated C-terminally tagged DadhA encoded on a separate pET28a vector. Anaerobic and aerobic expression and SDS-PAGE analysis revealed that the narG-dadhA fusion yielded the highest proportion of soluble protein, while the expression of DadhA appeared unchanged in the presence of a co-expressing chaperone DadhE (Figure 5.13). However, LC-MS analysis of culture supernatants and anaerobic lysate reactions incubated with dopamine indicated no construct yielded an active dopamine dehydroxylating enzyme.

**Table 5.1 Dadh constructs expressed on pET vectors.**

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Construct</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dadh-1</td>
<td>pET28a-dadhA_C-His₆</td>
<td></td>
</tr>
<tr>
<td>Dadh-3</td>
<td>pET8a-dadhA_noTATsignal_C-His₆</td>
<td>Endogenous dadhA TAT sequence removed</td>
</tr>
<tr>
<td>Dadh-4</td>
<td>pET28a-dadhA-noTATsignal_N-His₆</td>
<td>Endogenous dadhA TAT sequence removed</td>
</tr>
<tr>
<td>Dadh-5</td>
<td>pET8a-dadhA_narG-TAT_C-His₆</td>
<td>Endogenous dadhA TAT sequence replaced with E. coli narG TAT</td>
</tr>
<tr>
<td>7.1</td>
<td>pETDuet1_dadhE_N-His₆</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.13 Aerobic and anaerobic expression of DadhA in *E. coli* BL21 (DE3) in LB medium supplemented with nitrate, molybdate, and ferric ammonium citrate.

A) shows the aerobic expression and B) shows the anaerobic expression of the exact same constructs. First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. The right-most lane shows the empty vector control in which cells were co-transformed with pETDuet1 and pET28a vectors and exposed to IPTG. C = crude lysate fraction, S = soluble lysate fraction. Names of expression constructs are described in Table 5.1. The arrow indicates the location of DadhA.
Installation of the F0 cluster in the catalytic subunit is a prerequisite for proper bis-MGD cofactor incorporation (20, 66), suggesting that the lack of activity in the constructs described above could reflect a lack of proper F0 cluster assembly in DadhA. To increase iron-sulfur cluster production in the heterologous host, we transformed *E. coli* BL21 (DE3) with the pPH149 plasmid harboring the *E. coli* *iscSUA* genes that are necessary for iron-sulfur cluster assembly (100, 101). IscS is a master enzyme that delivers sulfur to partners involved in iron-sulfur cluster assembly and also assembles clusters on the scaffold protein IscU (102). The iron-binding protein IscA transfers the intact iron-sulfur cluster to other proteins (102). Based on the characterized roles of these proteins, we hypothesized that high levels of *iscSUA* expression enabled by pPPH149 would aid in DadhA metallocofactor assembly. Anaerobic and aerobic expression of various DadhA constructs in the *E. coli* BL21 (DE3) pPH149 background indicated similar expression patterns to those observed in the absence of the auxiliary plasmid, with the co-expression of the DadhE chaperone yielding a higher proportion of soluble enzyme. (Figure 5.14). We ensured that all co-transformed plasmids Although the lysates displayed different colors that correlated with the expression observed on the SDS-page gel (Figure 5.14), we did not observe any dopamine dehydroxylase activity by LC-MS.
Figure 5.14 Aerobic and anaerobic expression of DadhA in *E. coli* BL21 (DE3) harboring the pH149 plasmid in LB medium supplemented with nitrate, cysteine, molybdate, and ferric ammonium citrate.

A) shows the aerobic expression and B) shows the colors of the lysates from this expression. C) shows the anaerobic expression of the same constructs. First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. The two right-most lanes show the empty vector controls. C = crude lysate fraction, S = soluble lysate fraction. The arrow indicates the location of DadhA and the potential location of the DadhE chaperone. The shorthand names are described in Table 5.1.
The growth medium composition and expression protocol are two other factors that could influence metallocofactor assembly and Dadh activity. In our initial studies, we used LB medium supplemented with molybdate and iron, the two metals that Dadh is predicted to contain. Previously successful attempts at heterologous expression in *E. coli* (arsenate reductase and acetylene hydratase) used slightly different medium formulation and expression conditions (43, 99). We attempted expression of selected constructs in *E. coli* BL21 (DE3) (with and without the pH149 plasmid) adapting the protocols described in these two reports (43, 99). SDS-PAGE analysis indicated no major improvement in the proportion of soluble protein, although the expression of the N-terminally His\textsubscript{6}-tagged DadhE chaperone was clearly visible on the gel (Figure 5.15). However, no construct yielded active dopamine dehydroxylase, as assessed by LC-MS analysis of anaerobic cultures and lysates incubated with dopamine.
Figure 5.15 Anaerobic expression of selected DadhA constructs in *E. coli* BL21 (DE3) adapting previous protocols for acetylene hydratase and arsenate reductase expression. We adapted previously described protocols for heterologous expression of acetylene hydratase and arsenate reductase in *E. coli*. Selected constructs (shown in Table 5.1) were expressed in *E. coli* BL21 (DE3) with or without the pPH149 plasmid. A) shows the expression using conditions previously identified for acetylene hydratase. B) shows the expression using conditions previously identified for arsenate reductase. First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. C = crude lysate fraction, S = soluble lysate fraction. The arrow indicates the location of DadhA and the potential location of the DadhE chaperone. The shorthand names are described in Table 5.1.
Even though the protein appeared inactive independently of growth media, we attempted a preliminary purification of C-terminally His$_6$-tagged full length Dadh (on pET28a, named Dadh-1 in Table 5.1) co-expressed with the DadhE chaperone (on pETDuet1, named 7-1 in Table 5.1) and pH149 in *E. coli* BL21 (DE3). This construct had yielded the highest proportion of soluble protein in the small-scale expression assays described above (Figure 5.14). We attempted the purification from anaerobic and aerobic cultures (Figure 5.16). Under anaerobic conditions, the protein yield was so low that it was difficult to track the protein during the purification. In contrast, aerobic conditions yielded higher amounts of protein because *E. coli* BL21 (DE3) grew better under these conditions. However, in both cases a large amount of protein was present in the flow-through and many other proteins co-eluted with DadhA in the presence of higher imidazole concentrations. This could indicate poor binding to the column. Concentration of fractions harboring the band of interest produced a brown solution, suggesting potential presence of iron-sulfur clusters (Figure 5.16). However, this inactive protein was not analyzed further. Future work could assess the cofactor content of the inactive protein using ICP-MS and fluorescence spectroscopy (33, 35).
Figure 5.16 Attempted purification of C-terminally His₆-tagged full-length DadhA from E. coli BL21 (DE3).

Attempts to purify the C-terminally His₆-tagged Dadh construct (Dadh-1, described in Table 5.1). A) Shows the results from aerobic expression and B) shows the results from anaerobic expression. The picture insert shows the concentrated combined imidazole elution fractions, which had a brown-red color that could indicate presence of iron. In both gel images, the first lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards.
As discussed in section 5.1 of this Chapter, many bis-MGD enzymes are comprised of multiple subunits, and the genomic context of Dadh suggests it is likely a multi-subunit complex comprised of the catalytic bis-MGD-binding subunit, an electron shuttling ferredoxin, and a membrane anchor (Figures 5.9 and 5.12). Whereas the membrane anchor improves enzyme stability, the catalytic subunit and the partner ferredoxin are generally essential for activity (99, 103). We therefore attempted co-expression of the catalytic DadhA subunit together with different combinations of the other subunits, in the presence or absence of the DadhE chaperone (Table 5.2). We expressed these constructs alone or in combination in an E. coli BL21 (DE3) strain carrying a knock-out in iscR, the master regulator of iron-sulfur cluster assembly (104). Similar to the pH149 plasmid, we predicted that deletion of the repressor iscR should increase production of proteins involved in iron-sulfur cluster assembly. Anaerobic expression using the arsenate reductase and acetylene hydratase expression conditions described above revealed varying levels of protein expression. DadhA appeared mostly insoluble under these conditions, and in some cases expression levels were difficult to distinguish from background using SDS-PAGE (the only method used to assess expression). There was no clear expression of the ferredoxin or the anchor subunits. However, the chaperone DadhE was clearly expressed (Figure 5.17). Anaerobic lysates and cultures did not have dopamine dehydroxylase activity, again indicating inactive enzymes. Notably, DadhE appeared to be consistently expressed at higher levels in the presence of pH149 or in the ΔiscR background (Figures 5.15 and 5.17).

Our use of the pH149 plasmid and the ΔiscR E. coli strain attempted to boost iron-sulfur cluster incorporation and facilitate proper F0 cluster assembly in Dadh. We next attempted to instead increase synthesis of the bis-MGD cofactor. In our expression experiments, we could detect evidence for DMSO respiration in E. coli BL21 (DE3) (indicated by the formation of the volatile,
odorous product dimethylsulfide) which suggests that its endogenous bis-MGD-dependent enzyme DMSO reductase was maturing properly. Even so, it is possible that the bis-MGD cofactor was produced in insufficient amounts to become incorporated in the heterologously expressed Dadh. To boost bis-MGD cofactor formation, we co-expressed DadhA together with an auxiliary plasmid harboring the full *E. coli* operon expressing the *moa* biosynthetic genes. Expression of these genes should enable higher production of the MPT pterin intermediate, which could be further processed into bis-MGD (Figure 5.8). We received this plasmid (named pTPR1) from the laboratory of Dr. Tracy Palmer at Newcastle University. However, co-expression of the *moa* plasmid with the full three-gene *dadh* operon and chaperone on pETDuet1 yielded no obvious DadhA expression (Figure 5.17), despite the chaperone being clearly expressed. Again, LC-MS analysis of cultures and anaerobic lysates incubated with dopamine indicated no production of active protein.

All the constructs until this point had been in expressed in an *E. coli* BL21 host strain. Even though acetylene hydratase was successfully expressed in an *E. coli* BL21-derived strain (44), recent work has indicated that this strain harbors major metabolic deficiencies in metal import, anaerobic respiration, and bis-MGD-dependent reactions (105). For example, this strain lacks the genes encoding the high-affinity molybdate transport system and the molybdate-responsive transcriptional regulator ModE. It also lacks bis-MGD-dependent nitrate reductase and formate dehydrogenase activity and has a nonsense mutation in the *fnr* gene, a global regulator of anaerobic growth (105). This could explain why we consistently observed poor anaerobic growth of BL21 in the presence of DMSO and nitrate as the electron acceptors, which both require bis-MGD enzymes for metabolism, but not the electron acceptor fumarate, whose metabolism does not depend on molybdenum. This could also explain the lack of dopamine dehydroxylase activity of our
constructs. Thus, we expressed the full three-gene *dadhABC* operon or only the *dadhA* catalytic subunit and the partner ferredoxin *dadhB* in *E. coli* MG1655, a K12-derived *E. coli* strain that does not harbor deficiencies in anaerobic growth or molybdenum uptake (Table 5.2). We received this strain from the laboratory of Dr. Tracy Palmer at Newcastle University. She suggested this change in expression strain because *E. coli* BL21 had never been a suitable host for bis-MGD enzyme expression in her laboratory’s hands.

*E. coli* MG1655 is not lysogenized, meaning it cannot express plasmids driven by the strong T7 promoter present in vectors such as pETDuet1, pET28a, and pCOLA. Instead, we used the pTrcHis2A plasmid which uses the significantly weaker pTrc promoter to drive expression. The full three-gene construct yielded clear, soluble expression of a protein with the same size as the DadhA catalytic subunit (Figure 5.17). The expression was dependent on IPTG addition (Figure 5.18). We observed no clear expression of the other subunits using this construct. Moreover, we did not observe DadhA expression from the construct harboring only the DadhA catalytic subunit and its partner ferredoxin DadhB (Figure 5.17). However, all constructs were inactive, as assessed by LC-MS analysis of culture supernatants and anaerobic lysates incubated with dopamine. Co-expression of full three-gene *dadhABC* construct (on pTrcHis2A) with the pTPR1 moa gene plasmid did not change this result (Figure 5.19). Additionally, we expressed the pTrcHis2A-DadhABC construct together with the DadhE chaperone (on the pEXT22 plasmid) in an *E. coli* K-12-derived ΔiscR strain received from the *E. coli* KEIO collection. Surprisingly, however, we did not observe DadhA expression in this strain (Figure 5.19), and we also observed no dopamine dehydroxylase activity.
### Table 5.2 Strategy for Dadh multiple gene expression in E. coli.

<table>
<thead>
<tr>
<th>Strain name</th>
<th>Plasmid(s) transformed</th>
<th>E. coli expression strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pETDuet1-dadhC_N-His$_6$ _dadhE_C-strep/pCOLA-dadhA_C-strep _dadhB_N-His$_6$</td>
<td>BL21 (DE3) ΔiscR</td>
</tr>
<tr>
<td>2</td>
<td>pETDuet1-dadhABC_C-strep _dadhE_N-His$_6$</td>
<td>BL21 (DE3) ΔiscR</td>
</tr>
<tr>
<td>3</td>
<td>pCOLA-dadhA_C-strep _dadhB_N-His$_6$/PetDuet-dadhE_N-His$_6$</td>
<td>BL21 (DE3) ΔiscR</td>
</tr>
<tr>
<td>4</td>
<td>pTrcHis2A-dadhABC_C-His$_6$</td>
<td>MG1655</td>
</tr>
<tr>
<td>5</td>
<td>pTrcHis2A-dadhAB_C-His$_6$</td>
<td>MG1655</td>
</tr>
<tr>
<td>6</td>
<td>pETDuet1-dadhABC_C-strep _dadhE_N-His$_6$/pTPR1-moa operon</td>
<td>BL21 (DE3) ΔiscR</td>
</tr>
</tbody>
</table>

**Figure 5.17 Multiple gene expression of Dadh in E. coli BL21 (DE3) ΔiscR and E. coli MG1655.**

First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. C = crude lysate fraction, S = soluble lysate fraction. The arrow indicates the location of DadhA and the potential location of the DadhE chaperone. The shorthand names are described in Table 5.2.
Figure 5.18 Expression of DadhABC on pTrcHis2A in *E. coli* MG1655.
First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. C = crude lysate fraction, S = soluble lysate fraction. Subsequent two lanes indicate expression conditions where the inducer IPTG was added, while the rightmost lane represent an experiment where IPTG was not added. The arrow indicates the location of DadhA. This experiment revealed that the observed expression depends on IPTG. The pTrcHis2A-*dadhABC* plasmid is described in Table 5.2.
Table 5.3 Plasmids used for expression of DadhABC on pTrcHis2A in with and without auxiliary plasmids in *E. coli* MG1655 and *E. coli* K-12 ΔiscR.

<table>
<thead>
<tr>
<th>Shorthand</th>
<th>Plasmid</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTrc-1</td>
<td>pTrcHis2A-dadhABC_C-His6</td>
<td></td>
</tr>
<tr>
<td>pEXT22 chap</td>
<td>pEXT22-dadhE</td>
<td></td>
</tr>
<tr>
<td>pTPR1</td>
<td>pTPR1-moa_operon</td>
<td>Harbors <em>E. coli</em> moa genes for increased bis-MGD biosynthesis</td>
</tr>
</tbody>
</table>

Figure 5.19 Expression of DadhABC on pTrcHis2A in with and without auxiliary plasmids in *E. coli* MG1655 and *E. coli* K-12 ΔiscR.

First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. C = crude lysate fraction, S = soluble lysate fraction. The pTPR1 plasmid harbors the *moa* genes involved in bis-MGD biosynthesis. The pEXT22 plasmid harbors the chaperone DadhE, here abbreviated as “chap”. The plasmids are described in Table 5.3.
We also attempted expression in *E. coli* MG1655 (DE3), a lysogenized strain that would allow gene expression driven by the strong T7 promoter. We received this strain from the laboratory of Dr. Kristala Prather at MIT (106). We transformed this strain with the full Dadh operon pTrcHis2A construct as well as a pETDuet1 vector harboring N-terminally His-tagged Dadh chaperone and the full Dadh operon in the other multiple cloning site (Table 5.4). Despite apparent expression in this strain (Figure 5.20), we did not detect any dopamine dehydroxylase activity in anaerobic cultures or lysates.

Table 5.4 Plasmids used for multiple gene expression in *E. coli* MG1655 (DE3).

<table>
<thead>
<tr>
<th>Shorthand</th>
<th>Plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>D7XY</td>
<td>pETDuet1-dadhABC_C-strep_dadhE_N-His6</td>
</tr>
<tr>
<td>pTrc-1</td>
<td>pTrcHis2A-dadhABC_C-His6</td>
</tr>
</tbody>
</table>

Figure 5.20 Expression of DadhABCE on pETDuet1 in *E. coli* MG1655(DE3).
The plasmids were expressed under both anaerobic and aerobic conditions. First lane from the left shows the ladder, which is the Precision Plus Protein™ All Blue Standards. C = crude lysate fraction, S = soluble lysate fraction. The red arrows indicate the presence of DadhA (115 kDa) and DadhE (abbreviated as chap, 28 kDa) The plasmids are described in Tables 5.3 and 5.4.
Finally, it is possible that proteins beyond DadhABC and DadhE are required for proper dopamine dehydroxylase activity. For example, the ferredoxins DadhFGH and the hypothetical protein DadhD are upregulated in response to dopamine along with DadhABC and DadhE, suggesting potential involvement in enzyme maturation or activity. To test this hypothesis, we first cloned the *dadhABC* full three-gene operon into pTrcHis2A and in parallel cloned the second operon harboring *dadhEFGH* into the separate plasmid pBAD18K (Table 5.5). These were then successfully co-transformed into *E. coli* MG1655 but we have not yet had the opportunity to attempt protein expression experiments. The cloning and transformation were performed by Dr. Chip Le in our laboratory.

**Table 5.5 Plasmids yet to be co-expressed in *E. coli* MG1655.**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTrcHis2A-<em>dadhABC</em>&lt;sub&gt;C-H&lt;/sub&gt;6&lt;sub&gt;6&lt;/sub&gt;</td>
<td>Fully cloned, yet to be expressed</td>
</tr>
<tr>
<td>pBAD18K-<em>dadhEFGH</em></td>
<td>Fully cloned, yet to be expressed</td>
</tr>
</tbody>
</table>

**5.2.4 Expression of Dadh in *R. erythropolis* and *E. faecalis***

Bis-MGD enzyme assembly is a complex process that relies on interacting cellular systems that may only be present in the native host. Our failure to express active, soluble Dadh *E. coli* indicates necessary factors for proper Dadh assembly and activity were likely missing. The gram-negative proteobacterium *E. coli* is distantly related to the Dadh-encoding, gram-positive gut Actinobacterium *E. lenta*. In addition to the phylogenetic distance, the difference in membrane architecture between gram-positive and gram-negative bacteria could have implications for predicted TAT-exported and membrane-anchored proteins such as Dadh. This prompted us to attempt heterologous expression in more closely related hosts. We first explored expression in *Rhodococcus erythropolis*, a gram-positive soil Actinobacterium more closely related to *E. lenta* than *E. coli*. Notably, *R. erythropolis* L-88 was recently used to successfully express the *E. lenta*
enzyme Cgr2 that reduces the cardiac drug digoxin (97). We adapted this previously reported Cgr expression system to clone dadhABC into the thiostrepton-inducible pTipQC-2 shuttle vector (Table 5.6). We expressed the protein aerobically in LB in the presence of either added nitrate or DMSO to stimulate bis-MGD enzyme production. However, SDS-page revealed no obvious expression of any Dadh subunit. Consistent with this, LC-MS analysis of cultures and lysates incubated with dopamine revealed no m-tyramine production. This organism is an aerobe, preventing protein expression under anaerobic conditions.

Use of an obligate aerobe like R. erythropolis as an expression host, presents potential problems for the activity of the highly oxygen-sensitive Dadh enzyme. We therefore attempted Dadh expression in a gram-positive facultative anaerobe, selecting E. faecalis because of its genetic tractability and well-developed genetic toolkit. We cloned dadhABC into the shuttle vector pTRKH2 (107) and then transformed it into E. faecalis OG1RF. This facultatively anaerobic species harbors molybdenum-dependent enzymes (108), and we selected the specific strain because it has only minimal intrinsic antibiotic resistance (109). However, anaerobic expression produced no visible Dadh expression, as assessed by SDS-PAGE (Table 5.6). Consistent with this, LC-MS analysis of cultures and lysates incubated with dopamine revealed no m-tyramine production.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Expressed in</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTipQC-2-dadhABC</td>
<td>R. erythropolis L-88</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pTRKH2-dadhABC</td>
<td>E. faecalis OG1RF</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
</tbody>
</table>
5.2.5 Expression of additional catechol dehydroxylases in *E. coli* and *Shewanella* sp. ANA-3

Difficulties obtaining active, soluble enzyme can sometimes be overcome by expressing closely related homologs from either the same or different organisms (110). Having failed to obtain soluble, active Dadh from diverse plasmids and heterologous hosts, we next attempted expression of related catechol dehydroxylase homologs. We initially targeted the *E. lenta* hydrocaffeic acid dehydroxylase (Hcdh) whose biochemical activities we had confirmed through native purification (see Chapter 4). We also selected two putative catechol dehydroxylases from the gut Actinobacterium *Gordonibacter*: DOPAC dehydroxylase Dodh and the closely related catechol lignan dehydroxylase Cldh (111). Finally, we selected the dehydroxylase homolog (Tardh) involved in 3,5-dihydroxybenzoate (3,5-DHB) metabolism in the soil proteobacterium *T. aromatica*. Although the precise function of this enzyme remains unknown, it likely metabolizes one or both of the catecholic intermediates in the 3,5-DHB degradation pathway (112-114). The phylogenetic analysis of this sequence and a discussion of its potential function were provided in Chapter 4. Similar to dadh, hcdh is present in a three-gene operon consisting of the bis-MGD catalytic subunit (*hcdhA*), an electron-shuttling ferredoxin (*hcdhB*), and a membrane anchor (*hcdhC*) (Figure 5.21). In contrast, as described in Chapter 4, the *Gordonibacter* dehydroxylases Dodh and Cldh are predicted to be smaller, free-standing enzymes consisting of a catalytic bis-MGD subunit (*dodhA* or *cldhA*) and a small ferredoxin harboring a single 4Fe-4S cluster (*dodhB* or *cldhB*) (Figure 5.21). Tardh is a free-standing protein predicted to only comprise a catalytic bis-MGD subunit and no auxiliary proteins. Taken together, these sequences represent both a diversity in phylogenetic origin and predicted subunit compositions, which could increase the chance of obtaining pure, active protein. Our analysis of the predicted subunit compositions informed our cloning and expression strategy, as detailed below.
Figure 5.21 Predicted subunit composition of Dadh homologs selected for heterologous expression.
The indicated kb marker applies to all sequences.

We cloned all predicted subunits of dodh, cldh, tardh, and hcdh into the broad-host range, arabinose-inducible vector pBAD18K previously used to express the respiratory arsenate reductase (33). We also included the previously expressed dadh for comparison and as a control. This strategy enabled expression of a single plasmid across multiple heterologous hosts, including in E. coli BL21 and MG1655, and the related gram-negative proteobacterium Shewanella sp. ANA-3 (Table 5.7). We did not incorporate any type of affinity tag in these constructs because we were preliminary interested in gain-of-function in the heterologous host. Thus, we used gain of function assays to screen for activity. We did not observe any obvious anaerobic or aerobic expression of any of the subunits of any of the constructs by SDS-PAGE (Table 5.7). Furthermore, we did not observe any DOPAC, dopamine, or hydrocaffeic acid dehydroxylation activity by LC-
MS analysis of cultures and anaerobic lysates incubated with substrate (Table 5.7). We did not evaluate catechol lignan dehydroxylase activity due to the unavailability of the substrate dmSECO. Additionally, to evaluate Tardh activity, we used hydroxyhydroquinone as the substrate. It is possible that this enzyme acts on 2,3,5-trihydroxybenzoate (112, 113), but we did not evaluate this substrate due to the commercial unavailability of the compound. The cloning and expression of these constructs was performed by Sina Kiamehr in our laboratory.

We were surprised by the lack of expression of any of the genes under these conditions, especially considering our successful attempts at expressing dadh on different plasmids as described in the preceding section. Therefore, we cloned dodhAB and cldhAB into the pTrcHis2A plasmid that yielded successful expression of DadhA in E. coli MG1655 (Figure 5.19). However, we did not observe any visible expression by SDS-PAGE, nor did we detect DOPAC dehydroxylase activity by LC-MS analysis of cultures and anaerobic lysates incubated with substrate (Table 5.7). This cloning and expression was performed by Sina Kiamehr in our laboratory. Taken together, these results indicate that additional factors are necessary for dehydroxylase activity in these heterologous hosts.
Table 5.7 Expression of dehydroxylase homologs in *Shewanella sp. ANA-3* and *E. coli*.
Expression was assessed by SDS-PAGE. Activity was assessed by LC-MS analysis of lysate assays and cultures incubated with substrate.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Expressed in</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBAD18K-dadhABC</td>
<td>*Shewanella sp. ANA-3, E. coli BL21 (DE3) and MG1655</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pBAD18K-hcdhABC</td>
<td>*Shewanella sp. ANA-3, E. coli BL21 (DE3) and MG1655</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pBAD18K-dodhAB</td>
<td>*Shewanella sp. ANA-3, E. coli BL21 (DE3) and MG1655</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pBAD18K-cldhAB</td>
<td>*Shewanella sp. ANA-3, E. coli BL21 (DE3) and MG1655</td>
<td>No obvious expression (activity not assessed)</td>
</tr>
<tr>
<td>pBAD18K-tardh</td>
<td>*Shewanella sp. ANA-3, E. coli BL21 (DE3) and MG1655</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pTrcHis2A-dodhABCC-His6</td>
<td>*E. coli MG1655</td>
<td>No dehydroxylase activity or obvious expression</td>
</tr>
<tr>
<td>pTrcHis2A-cldhABC-His6</td>
<td>*E. coli MG1655</td>
<td>No obvious expression (activity not assessed)</td>
</tr>
</tbody>
</table>

As described in Chapter 4, *Gordonibacter* enzymes such as Dodh and Cldh are the dehydroxylases most closely related to acetylene hydratase, both in terms of amino acid sequence and predicted subunit composition. As described in Section 5.1 of this Chapter, the acetylene hydratase was previously successfully expressed in *E. coli* (43). In our final expression effort, we precisely followed the strategy and conditions used in the acetylene hydratase expression. We used Dodh as the model sequence, creating various constructs containing either *dadhA* alone or in combination with its ferredoxin partner *dadhB* on the pET vectors that were used in the previous study and that yielded soluble Dadh protein in our work (Figures 5.13 to 5.15). To some constructs, we also added the *E. coli* NarG TAT signal sequence to facilitate recognition by endogenous *E. coli* machinery and proper Dodh maturation (43). We cloned a total of 5 constructs and transformed these into *E. coli* Rosetta (DE3) and *E. coli* BL21 (DE3) (Table 5.8). *E. coli* Rosetta (DE3), the strain that was previously used in the acetylene hydratase expression, harbors an additional
plasmid encoding for rare tRNA codons, which can enable higher protein yield and purity (115). These constructs are due to be expressed in a minimal mineral medium (43), but we have not yet had the opportunity to evaluate expression and activity. The cloning was performed by Dr. Chip Le, Sina Kiamehr, and Minwoo Bae in our laboratory.

**Table 5.8 Expression of Dodh in E. coli.**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Comments</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET28a-dodhAB_C-His6</td>
<td></td>
<td>Fully cloned, yet to be expressed</td>
</tr>
<tr>
<td>pETDuet1-dodhA_C-His6</td>
<td></td>
<td>Fully cloned, yet to be expressed</td>
</tr>
<tr>
<td>pETDuet1-dodhA_C-His6_dodhB</td>
<td></td>
<td>Fully cloned, yet to be expressed</td>
</tr>
<tr>
<td>pET28a-narG-dodhAB_C-His6</td>
<td>E. coli narG TAT sequence added to N-terminus of dodhA</td>
<td>Fully cloned, yet to be expressed</td>
</tr>
<tr>
<td>pETDuet1-narG_dodhA_C-His6_dodhB</td>
<td>E. coli narG TAT sequence added to N-terminus of dodhA</td>
<td>Fully cloned, yet to be expressed</td>
</tr>
</tbody>
</table>
5.3 Discussion

Catechol dehydroxylases represent a new class of bis-MGD enzymes that are widely distributed among diverse microbes. However, the function of a majority of these enzymes remains unknown. These enzymes likely perform important ecological roles and participate in both primary and secondary microbial metabolism \((112, 116, 117)\). Furthermore, catechol dehydroxylases are chemically intriguing because of their ability to perform chemistry that is not possible using current synthetic methods and has been rarely described in enzymology. We have proposed that catechol dehydroxylases utilize a unique reaction mechanism to accomplish the chemically challenging aromatic ring reduction (Figure 5.7). This mechanism differs from that of the only other previously characterized aromatic dehydroxylase, 4-HBCR, which uses a radical Birch-like mechanism to remove the para-hydroxyl from 4-hydroxybenzoyl CoA. Our proposed mechanism is consistent with the Dadh and Hcdh substrate scope (described in Chapter 4) as well as the high (>95%) incorporation of a single deuterium into the product (Figures 5.10 and 5.11). Characterizing the mechanism of catechol dehydroxylation in greater detail will provide fundamental insight into enzymology, chemistry, and microbiology alike.

To facilitate discovery of new dehydroxylases as well as further biochemical, mechanistic, and structural characterization, we attempted heterologous expression of the dopamine dehydroxylase Dadh and related homologs. Screening expression of diverse Dadh constructs across distinct heterologous hosts, we successfully identified conditions to express soluble DadhA. However, all constructs produced inactive protein. Additionally, we expressed the dehydroxylase homologs Dodh, Cldh, Hcdh, and Tardh in \(E.\ coli\) and the related gram-negative proteobacterium \(Shewanella\) sp. ANA-3. We did not observe clear expression or activity of any of these constructs. Taken together, these results suggest that additional factors are required for heterologous
expression of soluble, active dehydroxylases, and set the stage for further cloning, expression, and characterization of catechol dehydroxylase family members.

Multiple factors could explain our failure to obtain active, soluble protein, including improper catechol dehydroxylase maturation. *E. coli* synthesizes the bis-MGD cofactor under the expression conditions used in our study (23, 33, 43) and we also found that overexpression of the *moa* biosynthetic genes did not yield active Dadh. Thus, in cases where we observed expression by SDS-PAGE, lack of activity likely derives from improper cofactor installation rather than a lack of cofactor biosynthesis. Future efforts could involve purifying Dadh and its homologs expressed in *E. coli* and *Shewanella sp.* ANA-3 and measure the metal and cofactor content of the inactive dehydroxylases using ICP-MS, UV-VIS spectroscopy, and fluorescence spectroscopy (33, 35, 97). This would highlight whether the inactive enzyme lacks either the bis-MGD cofactor or the iron-sulfur cluster, or both. It is also possible that heterologously expressed proteins harbor a terminal non-amino that differs from that in the native protein. To explore the potential presence of a sulfido ligand, the activity of natively purified dehydroxylases could evaluated upon exposure to cyanide, an anion that covalently binds the molybdenum atom and replaces the sulfido ligand, thus inactivating the enzyme (97).

Future purification efforts would be enabled by the addition of an affinity tag to the catalytic subunit, which was lacking in many of our constructs. Installation of an affinity tag would also allow for Western Blot analysis to assess expression. It is possible that some proteins that appeared to not be expressed from SDS-PAGE analysis would be visible by Western Blot due to a higher signal-to-noise ratio. Finally, future efforts could also assess the extent to which the dehydroxylase peptide is processed by *E. coli* pathways involved in bis-MGD enzyme maturation, including the TAT machinery. Our experiments with Dadh revealed that the TAT signal sequence
was important for expression, suggesting potential interactions with the *E. coli* machinery. To explore this possibility, future efforts could use fractionation to assess the cellular localization of the dehydroxylase (periplasm, membrane, cytoplasm). Additionally, proteomics could be used to assess whether the N-terminal TAT signal sequence has been cleaved off in the heterologously expressed dehydroxylase as was observed with the natively purified Hcdh protein described in Chapter 4.

In addition to assessing the cofactor content and processing of heterologously expressed dehydroxylases, future efforts should consider co-expression with additional genes from the native host. The lack of dopamine dehydroxylase activity in co-expression experiments with Dadh and its chaperone partner DadhE suggests that additional factors beyond DadhABCE and those available in *E. coli* could be required for dopamine dehydroxylase activity. Our RNA-sequencing experiments described in Chapter 4 suggested that each catechol induces the expression of not just a bis-MGD enzyme, but also of many genes in the surrounding genomic context. For example, dopamine induced the expression of the three-gene *dadhABC* operon as well as a separate operon harboring three distinct ferredoxins, the chaperone, and a protein of unknown function (Figure 5.5). We made similar observations with *hcdh* in *E. lenta* and *dodh* in *G. pamelaeae* 3C (See Chapter 4 for RNA-sequencing data). Co-expression of the gene(s) encoding the bis-MGD enzyme with surrounding genes upregulated in response to the substrate may improve enzyme maturation and activity.

Future efforts should also consider the use of alternative expression hosts. We have attempted expression across different heterologous hosts, but none of the strains that we used harbor catechol dehydroxylase homologs. *Streptomyces* species represent potentially promising hosts for dehydroxylase expression. Like human gut *Gordonibacter* and *Eggerthella*, these are
members of the Actinobacteria phylum that encode dehydroxylase homologs (as described in Chapter 4) (116, 117). Streptomyces strains have also been used for heterologous expression (118). For instance, the catechol dehydroxylase homolog involved in CC-1065 biosynthesis was successfully expressed on a plasmid in Streptomyces zelensis lacking the dehydroxylase, restoring the full biosynthetic pathway (116). The industrially important anaerobe Clostridium ljungdahlii represents another promising host, as this strain harbors a catechol dehydroxylase homolog (see Chapter 4) and is frequently used in metabolic engineering (119, 120).

Finally, future work could involve overexpression in the native host. While such an approach is not yet plausible in the human gut Gordonibacter and Eggerthella due to a lack of genetic tools, it should be possible in T. aromatica. Genetic tools and plasmids are available for this soil Proteobacterium, which encodes for the dehydroxylase homolog Tardh involved in 3,5-DHB degradation (112). Affinity-tagged Tardh could be cloned into a plasmid for expression under a strong promoter. This construct could then be expressed in T. aromatica grown with 3,5-DHB as the sole carbon source. The dehydroxylase homolog is necessary for robust growth under this condition, indicating that the necessary machinery for bis-MGD enzyme maturation is present (112). Finally, future efforts should consider expression in genetically tractable human gut anaerobes that encode for other bis-MGD enzymes (121-125).

Beyond heterologous expression, future efforts could consider in vitro reconstitution. This novel strategy was recently pioneered by Leimkuhler and co-workers in studies of the E. coli TMAO reductase (35). In their study, the authors first natively purified the Rhodobacter capsulatus molybdenum-dependent xanthine dehydrogenase, which harbors a metal-bound pyranopterin cofactor. Boiled R. capsulatus xanthine dehydrogenase, which released the cofactor upon denaturation, was then mixed with apo TMAO reductase purified from E. coli, the bis-MGD
synthesizing enzyme MobA, the dedicated TMAO bis-MGD insertion chaperone TorD, GTP, and magnesium under anaerobic conditions. This enabled MobA-dependent bis-MGD formation from the xanthine dehydrogenase-derived pterin and subsequent TorD-assisted cofactor insertion into the TMAO reductase to generate active holo enzyme (35). While this strategy offers unique control of the bis-MGD cofactor insertion process and circumvents the need for complex cellular machinery, it is unclear how it would work for more complex enzymes that, unlike TorD, harbor N-terminal 4Fe-4S clusters or function in complex with auxiliary proteins. One possible avenue includes expressing a gut Actinobacterial dehydroxylase such as Dadh or Dodh in E. coli and then purifying the inactive protein. This purified inactive protein could then be subjected to the same reconstitution process as the TMAO reductase, potentially installing the bis-MGD cofactor to generate active enzyme (35). If the 4Fe-4S clusters have not been properly installed inside the heterologous host, these could be reconstituted into the purified protein as well (97).

Finally, it is also worth considering native purification, the strategy most commonly used for bis-MGD enzyme purification. We confirmed the biochemical functions of Dadh and Hcdh using native purification from E. lenta, highlighting the viability of this strategy for verifying enzyme activity (see Chapters 3 and 4). However, moving beyond confirmation to a complete mechanistic, structural, and biochemical characterization will require significant improvements in the enzyme preparation, including increased yield and purity. Unlike in many other native purification experiments, we grew E. lenta in a rich medium where the substrate was not required for robust growth. Future efforts could modulate the medium composition to increase the relative expression of the catechol dehydroxylase. For example, we observed that catechols promote growth of gut Actinobacteria in a minimal growth medium lacking electron acceptors. While the overall biomass yield was low in this medium, the organism relies on the catechol dehydroxylase...
for robust growth, suggesting it may be highly expressed under these conditions (see Chapter 4 for experimental results). This minimal medium could therefore be optimized to increase biomass while maintaining the catechol-dependent growth advantage, potentially increasing the relative purity and yield of natively purified proteins. In summary, the work described here in this Chapter, along with work described in Chapter 4, has laid the foundation for further biochemical studies of catechol dehydroxylases.
5.4 Methods

5.4.1 General materials

The following chemicals were used in this study: Ferric ammonium citrate (Sigma-Aldrich, catalog# F5879-100G), sodium nitrate (Sigma-Aldrich, catalog# S5506-250G), Thiostrepton (Sigma-Aldrich, catalog# T8902-1G), sodium formate (Sigma-Aldrich, catalog# 71539-500G), glycine (Sigma-Aldrich, G8898-500G), DMSO (Sigma-Aldrich, catalog# D8418-50mL), dopamine (Sigma-Aldrich, catalog# PHR1090-1G), m-tyramine (Santa Cruz Biotechnology, catalog# sc-255257), isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich, catalog# I5502-1G), sodium molybdate (Sigma-Aldrich, catalog # 243655-100G), SIGMAFAST protease inhibitor tablets (Sigma-Aldrich, catalog#: S8830), benzyl viologen (Sigma-Aldrich, catalog# 271845-250mg), methyl viologen (Sigma-Aldrich, catalog# 856177-1g), diquat (Sigma-Aldrich, catalog# 45422-250mg), sodium dithionite (Sigma-Aldrich, catalog# 157953-5G), 3,4-dihydroxyphenylacetic acid (Millipore Sigma, catalog# 850217-1G), 3,4-dihydroxyhydrocinnamic acid (hydrocaffeic acid) (Millipore Sigma, catalog# 102601-10G), 1,2,4-benzentriol (Sigma-Aldrich, catalog# 173401-1G). Acetonitrile and methanol for LC-MS analyses were purchased as LC-MS grade solvent from Honeywell Burdick & Jackson or Sigma-Aldrich. Luria-Bertani (LB) medium was prepared from its basic components (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) or obtained from either EMD Millipore or Alfa Aesar. Terrific Broth (TB) was purchased from Alfa Aesar (catalog# AAH2682422) and supplemented with 4 mL/L glycerol and 10 mM MgSO₄. Gibson assembly master mix was purchased from NEB (catalog# E2611S). Q5 high-fidelity 2x PCR master mix was also purchased from NEB (catalog# M0492S).
5.4.2 General methods

All genomic DNA (gDNA) was extracted from bacterial cultures using the DNeasy UltraClean Microbial Kit (Qiagen, catalog #: 12224-50) according to the manufacturer’s protocol.

All anaerobic culturing work was performed in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 10% hydrogen, 10% carbon dioxide, and nitrogen as the balance, unless otherwise noted. All anaerobic lysate work and biochemical experiments were performed in an anaerobic chamber (Coy Laboratory Products) situated in a cold room at 4 °C under an atmosphere of 10% hydrogen and nitrogen as the balance.

Gut Actinobacterial strains were grown in BHI medium containing 1% arginine (w/v) to obtain pellets for gDNA extraction. *T. aromatica* AR-1 gDNA was obtained from the laboratory of Dr. Silvia Marques (CSIC, Estacion Experimental del Zaidin, Granada, Spain).

All PCR reactions were performed following the manufacturer’s (NEB) protocol. Briefly, each 20 µL PCR reaction from genomic DNA included: 10 µL Q5 master mix, 1 µL forward primer (10 µM final concentration), 1 µL reverse primer (10 µM final concentration), 2 µL gDNA (50-100 ng), and 6 µL water. Each 20 µL PCR reaction from involving a plasmid included: 10 µL Q5 master mix, 1 µL forward primer (10 µM final concentration), 1 µL reverse primer (10 µM final concentration), 2 µL gDNA (50-100 ng), and 6 µL water. The primer annealing temperature was selected based on the NEB annealing temperature calculator (http://tmccalculator.neb.com). We followed the generally recommended Q5 protocol for the PCR reactions: PCR was performed on a CFX96 Thermocycler (Bio-Rad), using the following program: initial denaturation at 98 °C for 30 seconds, 34 cycles of 98 °C for 10 seconds, selected annealing temperature for 30 seconds, 72 °C for 30 seconds/kb of amplified DNA. The program ended with a final extension at 72 °C for 5 mins.
All cloning work was performed following a general protocol, unless otherwise indicated. All PCR amplicons were purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, catalog# 28-9034-70) or the Zymoclean Gel DNA Recovery Kit (Zymo research, catalog# D4001). For Gibson assembly-based cloning, purified PCR products were then subjected to Gibson assembly following the manufacturer’s (NEB) protocol. Briefly, 50 ng of vector was mixed with 3-fold molar excess of insert, 5 µL Gibson Assembly 2x Mastermix (New England Biolabs, catalog# E2611S), and MilliQ water to a final volume of 10 µL. The Gibson reactions were left at 50 °C for 30 minutes and 5 µL of the reaction was transformed into chemically-competent E. coli TOP10 using heat shock. Insert presence was confirmed by PCR and/or DNA sequencing.

All restriction digestion work was performed as follows. NcoI (catalog# R3193S), and XhoI (catalog# R0146s) were purchased from New England Biolabs. For restriction digestion, 500 ng plasmid or 50 ng PCR insert were mixed with 1 µL of each restriction enzyme and 4 µL 10x cutsmart buffer (New England Biolabs, catalog# B7200S), and MilliQ water to a final reaction volume of 40 µL. Restriction digestion reactions were left at 37 °C for 3 hours, followed by gel purification using the GFX PCR DNA and Gel Band Purification Kit (GE healthcare, catalog# 28-9034-70). Ligation of purified digested vectors and inserts was performed by Gibson Assembly following the general protocol described above.

Heat shock was used to transform chemically competent E. coli strains. We used electroporation to transform R. erythropolis, and E. faecalis. We used conjugation to transform Shewanella sp. ANA-3.
5.4.3 LC-MS methods

Method A: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801). The flow rate was 0.5 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 0% B isocratic, 2-9 min: 0-10% B, 9-11 min: 10-95% B, 11-15 min: 95% B isocratic, 15-18 min: 95-0% B, 18-21 min: 0% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of dopamine (precursor ion \(m/z = 154.3\), daughter ion \(m/z = 137.3\)), and tyramine (precursor ion \(m/z = 138.3\), daughter ion \(m/z = 121.3\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in positive MRM mode.

Method B: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min\(^{-1}\) using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of hydrocaffeic acid (precursor ion \(m/z = 181.2\), daughter ion \(m/z = 137.2\)), hydroxyphenylpropionic acid (precursor ion \(m/z = 165.1\), daughter ion \(m/z = 121.2\)), DOPAC (precursor ion \(m/z = 167.2\), daughter ion \(m/z = 123.2\)), and hydroxyphenylacetic acid (precursor ion \(m/z = 151.3\), daughter ion \(m/z = 107.3\)) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

Method C: Samples were analyzed using an Agilent technologies 6530 Accurate-Mass Q-TOF LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801).
The flow rate was 0.4 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 5% B isocratic, 2-25 min: 0-95% B, 25-30 min: 95% B isocratic, 30-40 min: 95-5% B. For the MS detection, the ESI mass spectra data were recorded in negative mode for a mass range of m/z 50 to 3000. A mass window of ±0.005 Da was used to extract the ion of [M-H].

Method D: Samples were analyzed using an Agilent technologies 6530 Accurate-Mass Q-TOF LC/MS and a Dikma Technologies Inspire Phenyl column (4.6 × 150 mm, 5 μm; catalog #81801). The flow rate was 0.4 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. The column temperature was maintained at room temperature. The following gradient was applied: 0-2 min: 5% B isocratic, 2-25 min: 0-95% B, 25-30 min: 95% B isocratic, 30-40 min: 95-5% B. For the MS detection, the ESI mass spectra data were recorded in positive mode for a mass range of m/z 50 to 3000. A mass window of ±0.005 Da was used to extract the ion of [M+H].

Method E: Samples were analyzed using an Agilent technologies 6410 Triple Quad LC/MS and a Thermo Scientific Acclaim Polar Advantage II column (3 μM, 120A, 2.1*150 mm, product #: 063187). The flow rate was 0.2 mL min⁻¹ using 0.1% formic acid in water as mobile phase A and methanol as mobile phase B. The following gradient was applied: 0-4 min: 50% B isocratic, 4-7 min: 50-99%, 7-9 min: 99-50%, 9-13 min: 50% B isocratic. For mass spectrometry, the source temperature was 300 °C, and the masses of singly deuterated hydrocaffeic acid (precursor ion m/z = 182.2, daughter ion m/z = 138.2), singly deuterated hydroxyphenylpropionic acid (precursor ion m/z = 166.1, daughter ion m/z = 122.2), doubly deuterated hydrocaffeic acid (precursor ion m/z = 183.2, daughter ion m/z = 139.2), doubly deuterated hydroxyphenylpropionic acid (precursor ion
m/z = 167.1, daughter ion m/z = 123.2) were monitored at a collision energy of 15 mV and fragmentor setting of 135 in negative MRM mode.

5.4.4 Lysate assays in deuterated buffer in *E. lenta* A2

Bacterial cultures were grown in Hungate tubes. All bacterial growth and lysate experiments were performed in an anaerobic chamber. Lysis and sample processing took place in an anaerobic chamber kept at 4 °C. *E. lenta* A2 was inoculated from a single colony into 10 mL of BHI liquid medium and grown for 48 hours at 37 °C to provide turbid starter cultures. These were diluted 1:100 in triplicate into 50 mL of BHI medium containing 1% arginine and 10 mM formate and 1 mM hydrocaffeic acid or vehicle. After 18 hours of anaerobic growth at 37 °C, cultures had reached an OD₆₀₀ of approximately 0.700 and were harvested by centrifugation. The bacterial pellets were resuspended anaerobically in 20 mL of cold, pre-reduced 50 mM Tris pH 8 in D₂O to wash the cells, followed by an additional round of centrifugation to pellet the washed cells. Cells were washed twice more in this buffer. The washed cells from each culture were then resuspended in 0.80 mL of lysis buffer (50 mM Tris pH 8 in D₂O containing 4 mg/mL SIGMAFAST protease inhibitor cocktail) and transferred to an Eppendorf tube. The cells were lysed using sonication in an anaerobic chamber. 50 µL of this lysate was transferred in triplicate to a 96 well plate (VWR, catalog# 82006-636). 1 µL of substrate was then added to each of the replicates at a final concentration of 0.5 mM. We also added methyl viologen (1 µL, 1 mM final concentration) and sodium dithionite (2 µL, 2 mM final concentration) to these reaction mixtures (the final reaction mixture volume was 100 µL). All reaction components were dissolved in deuterated water as well. These samples were incubated anaerobically at room temperature for 28 hours to allow for metabolism to proceed. Samples were then analyzed by LC-MS. To prepare samples for LC-MS, 20 µL of the culture supernatant was diluted 1:10 with 180 µL of methanol, followed by
centrifugation at 4000 rpm for 10 minutes to pellet particulates, salts, and proteins. 50 µL of the resulting supernatant was then transferred to the LC-MS 96-well plate and 5 µL of the supernatant was injected onto the instrument using Method B to detect hydrocaffeic acid and Method E for deuterated metabolites. Method D was used for high-resolution LC-MS.

### 5.4.5 Cloning of pET28a-dadhA_C-His₆

pET28a was digested with the enzymes NcoI and ZhoI following the protocols described above. *E. lenta* A2 gDNA was PCR-amplified with the primers 5’ TTAAGAAGGAGATATACCAGTGCACGTACAGCGTGTC 3’ and 5’ TGGTGGTGGTGGCTCCTCCCTCCCTCCTCGTAGGT 3’ to generate the *dadhA* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

### 5.4.6 Cloning of pET28a-dadhA_noTATsignal_C-His₆

pET28a was digested with the enzymes NcoI and XhoI following the protocols described above. *E. lenta* A2 gDNA was PCR-amplified with the primers 5’ TTAAGAAGGAGATATACCGAGACGTACAGCGTGTC 3’ and 5’ TGGTGGTGGTGGCTCCTCCCTCCCTCCTCGTAGGT 3’ to generate the *dadhA* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

### 5.4.7 Cloning of pET28a-dadhA_noTATsignal_N-His₆

pET28 was digested with the enzymes NcoI and XhoI following the protocols described above. *E. lenta* A2 gDNA was PCR-amplified with the primers 5’ GGTGCCGCGCAGCAGCATGACGTACAGCAGCGTGTC 3’ and 5’ TGGTGGTGGTGGCTCCTCCCTCCCTCCTCGTAGGT 3’ to generate the *dadhA* amplicon.
The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.8 Cloning of pET28a-dadhA_narG-TAT_C-His\(_6\)

pET28a was digested with the enzymes NcoI and XhoI following the protocols described above. *E. lenta* A2 gDNA was PCR-amplified with the primers 5’ ATGGGTAACCTGACCATGTC 3’ and 5’ TGGTGGTGATGCTCTCCCTCCCTCCTGATGTTG 3’ to generate the *dadhA* amplicon. *E. coli* MG1655 gDNA was amplified with the primers 5’ TTAAGAAGGAGATATACCATGAGTAAATTCCTGGACCG 3’ and 5’ GACATGGTCAGTTACCATATATCCACTCCCTCCAGTCAC 3’ to generate the *narG* amplicon. The gel-purified vector and the two inserts (*dadhA*, *narG* N-terminus) were assembled using Gibson Assembly following the protocol described above.

5.4.9 Cloning of pETDuet1_dadhE_N-His\(_6\)

pETDuet1 was PCR-amplified with the primers 5’ CTCGAGTCTGGTAAAGAAA 3’ and 5’ CGATATCCAAATTGAGATCTGC 3’ to linearize the plasmid. *E. lenta* A2 gDNA was amplified with the primers 5’ GCAGATCTCAATTGGATATCGATGTCCCACGATCAGGAAACC 3’ and 5’ GGTTCATTACCAGACTCGAGTGCAACCTCGACTTCGGT 3’ to generate the *dadhE* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.10 Cloning of pCOLA-dadhA_C-strep_dadhB_N-His\(_6\)

pCOLADuet was first PCR-amplified with the primers 5’ GACAAGCTTGGCGCCGCA 3’ and 5’ GCTCGAATTCCGGATCCTGGC 3’ to linearize the plasmid for insertion into the
Multiple cloning site carrying the N-His\textsubscript{6} tag. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ AGCCAGGATCCGAATTCCGACCATGACGAAGGCTATCATC 3’ and 5’ TGCGGCCGCAAGCTTGTTACTCATCACCTCTTCCCCA 3’ to generate the \textit{dadhB} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above, creating the pCOLA- \textit{dadhB}_N-His\textsubscript{6} vector. This was then PCR-amplified with the primers 5’ GGTACCCTCGAGTCTGGT 3’ and 5’ CGATATCCAAATTGAGATCTGCTC 3’ to linearize the plasmid for insertion into the multiple cloning site carrying the C-Strep tag. \textit{E. lenta} A2 gDNA was then amplified with the primers 5’ GCAGATCTCAATTGGATATCGATGGGTAACCTGACCATG 3’ and 5’ ACCAGACTCGAGGATACCTCCCTTCTCTTCTGATGGTAGG 3’ to generate the \textit{dadhA} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above. This created the final pCOLA-\textit{dadhA}_C-strep_\textit{dadhB}_C-His\textsubscript{6} vector.

\textbf{5.4.11 Cloning of pETDuet1-\textit{dadhC}_N-His\textsubscript{6}_\textit{dadhE}_C-strep}

pETDuet1 was first PCR-amplified with the primers 5’ TAATGCTTAAGTCAAGACAGAA 3’ and 5’ CACAGCCAGGATCCGAATTCCG 3’ to linearize the plasmid for insertion into the multiple cloning site carrying the N-His\textsubscript{6} tag. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ CACAGCCAGGATCCGAATTCCGGGAAACGCAGTGGCCTCTC 3’ and 5’ TTCTGTTCAGCTTAAGTCAATTAGTACATCAGGACCACGTGG 3’ to generate the \textit{dadhC} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above, creating the pETDuet1-\textit{dadhC}_N-His\textsubscript{6} vector. This was then PCR-amplified with the primers 5’ CTCGAGTCTGGTAAAGAAACC 3’ and 5’
CGATATCCAATTGAGATCTGC 3’ to linearize the plasmid for insertion into the multiple cloning site carrying the C-strep tag. *E. lenta* A2 gDNA was then amplified with the primers 5’ GCAGATCTCAATTGGATATCGATGTCACCACGATCAGGAAACC 3’ and 5’ GGTNTTCTTTACCAGACTCGAGTCATGCAACCTCGACTTCGGT 3’ to generate the *dadhE* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above. This created the final pETDuet1-*dadhC_N-His6_dadhE_C-strep* vector.

**5.4.12 Cloning of pETDuet1-*dadhABC_C-strep_dadhE_N-His6***

pETDuet1 was first PCR-amplified with the primers 5’ TAATGCTTAAGTCAACAGAA 3’ and CACAGCCAGGATCCGAATTG to linearize the plasmid for insertion into the multiple cloning site carrying the N-His6 tag. *E. lenta* A2 gDNA was amplified with the primers 5’ TTCTGTTCAGACTTAAGCATTATCATGCAACCTCGACTTCGGT 3’ and 5’ CACAGCCAGGATCCGAATTGATGTCACCACGATCAGGAAACC 3’ to generate the *dadhE* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above, creating the pETDuet1-*dadhE_N-His6* vector. This was then PCR-amplified with the primers 5’ TCTCGAGTCTGGTAAAGCAAACC 3’ and 5’ CGATATCCAATTGAGATCTGC 3’ to linearize the plasmid for insertion into the multiple cloning site carrying the C-strep tag. *E. lenta* A2 gDNA was then amplified with the primers 5’ GCAGATCTCAATTGGATATCGATGTCACCACGATCAGGAAACC 3’ and 5’ GGTNTTCTTTACCAGACTCGAGTTAGTACATGCAACCTCGACTTG 3’ to generate the *dadhABC* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly
following the protocol described above. This created the final pETDuet1-\textit{dadhABC}_C-strep\_\textit{dadhE}_N-His\_6 vector.

\textbf{5.4.13 Cloning of pTrcHis2A-\textit{dadhABC}_C-His\_6}

pTrcHis2A was PCR-amplified with the primers 5’ GAACAAAAACTCATCTCAGAA 3’ and 5’ CATGGTTTATTCCCTCTTATT 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ AATAAGGAGGAATAAACCATGATGGGTAACCTGACCATGTCA 3’ and 5’ TTCTGAGATGAGTTTTTTGTTCTTAGATCAGGACCACCCTGG 3’ to generate the \textit{dadhABC} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

\textbf{5.4.14 Cloning of pTrcHis2A-\textit{dadhAB}_C-His\_6}

pTrcHis2A was PCR-amplified with the primers 5’ GAACAAAAACTCATCTCAGAA 3’ and 5’ CATGGTTTATTCCCTCTTATT 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ AATAAGGAGGAATAAACCATGATGGGTAACCTGACCATGTCA 3’ and 5’ TTCTGAGATGAGTTTTTTGTTCTTAGATCAGGACCACCCTGG 3’ to generate the \textit{dadhAB} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

\textbf{5.4.15 Cloning of pEXT22-\textit{dadhE}}

pEXT22 was PCR-amplified with the primers 5’ CCTGCAGGCGATGCAAGCT 3’ and 5’ CCCGGGTACCGAGCTCGA 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ TCGAGCTCGGTACCCGGGATGTCCCACGATCAGGAA 3’ and 5’ CCTGCAGGCGATGCAAGCTTCATGCAACCTCGACTTC 3’ to generate the \textit{dadhE} amplicon.
The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.16 Cloning of pBAD18K-\textit{dadhDEFGH}

pBAD18K was PCR-amplified with the primers 5’ TTGGCTGTTTTGGCGGATGAG 3’ and 5’ GAATTCGCTAGCCCCAAAACGGG 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ CCCGTTTTTTTTGGGCTAGCGAATTCAGGAATAAAATGAGCGATGTTGATGTATT 3’ and 5’ CTCATCCGCCAAAAACAGCCAACGCAATAGCACATGACTGATCC 3’ to generate the \textit{dadhDEFGH} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.17 Cloning of pTipQC-2-\textit{dadhABC}

pTipQC-2 was PCR-amplified with the primers 5’ CGATCCAAGCTTATGCT 3’ and 5’ GCCCATATGTATATCTCCTTC 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ GAAGGAGATATACATATGGGCATGGGTAACCTGACCATG 3’ and 5’ AGATCTAAGCTTGGGATCGTTAGTACATCAGGACCACTGG 3’ to generate the \textit{dadhABC} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.18 Cloning of pTRKH2-\textit{dadhABC}

pTRKH2 was PCR-amplified with the primers 5’ CCTGAATGCGAATGCTA 3’ and 5’ AGCTGTTTCTGCTGAA 3’ to linearize the plasmid. \textit{E. lenta} A2 gDNA was amplified with the primers 5’ TTCACACAGGAAACAGCTATGGGTAACCTGACCATG 3’ and 5’ TAGCATTGGCCATTAGTTAGTACATGACCAGACC 3’ to generate the \textit{dadhABC}
amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.19 Cloning of pBAD18K-dadhABC

pBAD18K was PCR-amplified with the primers 5’ GAATTTCGCTAGCCAAAAAAACGG 3’ and 5’ AAGCTTGGCTTTTTGGCG 3’ to linearize the plasmid. *E. lenta* A2 gDNA was amplified with the primers 5’ CCCGTTTTTTGCTAGCAGAATTCAGGAGATATACCATGGGTAACCTGACCAGCTGTC 3’ and 5’ CCGCCAAAAACAGCCAAAGCTTTTAGTACTACAGGACCAGCTGGGT 3’ to generate the *dadhABC* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.20 Cloning of pBAD18K-hcdhABC

pBAD18K was PCR-amplified with the primers 5’ GAATTTCGCTAGCCAAAAAAACGG 3’ and 5’ AAGCTTGGCTTTTTGGCG 3’ to linearize the plasmid. *E. lenta* A2 gDNA was amplified with the primers 5’ CCCGTTTTTTTGGCTAGCAGAATTCAGGAGATATACCATGATCAACGAATACCA 3’ and 5’ CCGCCAAAAACAGCCAAAGCTTTTAGTACAAAGCGCAGATGCT 3’ to generate the *hcdhABC* amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.21 Cloning of pBAD18K-dodhAB

pBAD18K was PCR-amplified with the primers 5’ GAATTTCGCTAGCCAAAAAAACGG 3’ and 5’ AAGCTTGGCTTTTTGGCG 3’ to linearize the plasmid. *G. pamelaeae* 3C A2 gDNA was amplified with the primers 5’ CCCGTTTTTTGCTAGCAGAATTCAGGAGATATACCATGATCAACGAATACCAACCA
T 3’ and 5’ CCGCCAAAACAGCCAAGCTTTTACTTTGCGAAAAGCACCCTGATT 3’ to generate the dodhAB amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.22 Cloning of pBAD18K-cldhAB

pBAD18K was PCR-amplified with the primers 5’ GAATTCGCTAGCCCAAAAAAACGG 3’ and 5’ AAGCTTGGCTGTTTTGGCG 3’ to linearize the plasmid. G. pamelaeae 3C gDNA was amplified with the primers 5’ CCCGTTTTTTTGGGCTAGCGAATTCAGGAGATATACCATGGAGTACAACACGAGTA C 3’ and 5’ CCGCCAAAACAGCCAAGCTTTTACTTTGCGAAAAGCACCCTGATT 3’ to generate the cldhAB amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.23 Cloning of pBAD18K-tardh

pBAD18K was PCR-amplified with the primers 5’ GAATTCGCTAGCCCAAAAAAACGG 3’ and 5’ AAGCTTGGCTGTTTTGGCG 3’ to linearize the plasmid. T. aromatica AR-1 gDNA was amplified with the primers 5’ CCCGTTTTTTTGGGCTAGCGAATTCAGGAGATATACCATGAAATGCTGTTTTGGA G 3’ and 5’ CCGCCAAAACAGCCAAGCTTTTATTGGCAGGAGGTTGATCGTC 3’ to generate the tardh amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.24 Cloning of pTrcHis2A-dodhAB

pTrcHis2A was PCR-amplified with the primers 5’ GAACAAAAACTCATCTCAGAA 3’ and 5’ CATGGTTTTATTCCCTCATT 3’ to linearize the plasmid. G. pamelaeae 3C gDNA was amplified with the primers 5’
AATAAGGAGGAATAAACCATGATGATCAACGAATACACCAT 3’ and 5’
TTCTGAGATGAGTTTTGTTCCTTTTCGAAAAGCACCTGATT 3’ to generate the dodhAB amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.25 Cloning of pTrcHis2A-cldhAB

pTrcHis2A was PCR-amplified with the primers 5’ GAACAAAAACTCATCTCAGAA 3’ and 5’ CATGGTTTATTCCTCTTATT 3’ to linearize the plasmid. *G. pamelaeae* 3C gDNA was amplified with the primers 5’
AATAAGGAGGAATAAACCATGATGAGTGATACACCATGATAC 3’ and 5’
TTCTGAGATGAGTTTTTGTTCCTTCGAGACCTGATT 3’ to generate the cldhAB amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.26 Cloning of pET28a-dodhAB_C-His₆

pET28a was PCR-amplified with the primers 5’ CTCGAGCACCACCACCAC 3’ and 5’
GGTATATCTCCTCTAAAGTTAAACAAAA 3’ to linearize the plasmid. *G. pamelaeae* 3C gDNA was amplified with the primers 5’
CTCTAGAAAAATTTTTTGTTTAAGGGGAGATATACCATGATCAACGAATAC 3’ and 5’
AACCATG 3’ and 5’
GTTAGCAGCCGATCTCAGTGTTGGTGTGTTGGTCTCGAGCTTTCGAAAAGCACCTG 3’ to generate the dodhAB amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.
5.4.27 Cloning of pETDuet1-\textit{dodhA}\textsubscript{C}-\textit{His\textsubscript{6}}

\textit{pETDuet1} was PCR-amplified with the primers 5’ CTTAAGTCAACGAAAGTAATCG 3’ and 5’ GGTATATCTCCTCTTAAAGTTAA 3’ to linearize the plasmid. \textit{G. pamelaeae} 3C gDNA was amplified with the primers 5’ CTCTAGAAATAATTGTTAACGTTAAGAAGGAGATATACCATGATCAACGAATAC AACCATGATCTGTAC 3’ and 5’ CCGTGTACAATACGATTTACTTTCTGTCCGACTTAAAGTCAAGTCGTTGTGCTGATGGGGTGGT to generate the \textit{dodhA} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above.

5.4.28 Cloning of pETDuet1-\textit{dodhA}\textsubscript{C}-\textit{His\textsubscript{6}}-\textit{dodhB}

\textit{pETDuet1-dodhA}\textsubscript{C}-\textit{His\textsubscript{6}} was PCR-amplified with the primers 5’ CACATGGACTCGTCTACTAG 3’ and 5’ ATGTATATCTCCTCTTATACCATTAAC 3’ to linearize the plasmid. \textit{G. pamelaeae} 3C gDNA was amplified with the primers 5’ CCATCTTAGTAGTATTAGTTAAGTAAGAAGGAGATATACCATGACTACACCCGTT TCG 3’ and 5’ GGTTAATAGCTGCAGTATACGAGGCCGATCTTCTTGGCAAAAGCACCCTG 3’ to generate the \textit{dodhB} amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above. This created the full \textit{pETDuet1-dodhA}\textsubscript{C}-\textit{His\textsubscript{6}}-\textit{dodhB} vector.
5.4.29 Cloning of pET28a-narG-dodhAB_C-His6

pET28a-dodhAB_C-His6 was PCR-amplified with the primers 5’ ATCAACGAATACAAACCAT 3’ and 5’ TATCTCCTCTTTAAAGTTAAAC 3’ to linearize the plasmid. G. pamelaeae 3C gDNA was amplified with the primers 5’ CTAGAAATAATTTTGTTTAACTTTAAGGAAGGATAATGAGTAAATTCTGGAC 3’ and 5’ GTACAGATCATGGTTGTATTCTTGATGTATATATATCCATCCTCCCAGTC 3’ to generate the narG amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above. This created the full pET28a-narG-dodhAB_C-His6 vector.

5.4.30 Cloning of pETDuet1-narG_dodhA_C-His6_dodhB

pETDuet1-dodhA_C-His6_dodhB vector was PCR-amplified with the primers ATCAACGAATACAAACCATG and TATCTCCTCTTTAAAGTTAAAC to linearize the plasmid. E. coli MG1655 gDNA was amplified with the primers 5’ GAAATAATTTTGAACCTTTAAGGAAGGATAATGAGTAAATTCTGGAC 3’ and 5’ GTTTTGTCAGATCATGGTTGTATTCTTGATGTATATATATCCATCCTCCCAGTC 3’ to generate the narG amplicon. The gel-purified vector and insert were assembled using Gibson Assembly following the protocol described above. This created the full pETDuet1-narG_dodhA_C-His6_dodhB vector.

5.4.31 Protein expression in E. coli

The protein expression trials described below are grouped into individual experiments as presented in the results section.
5.4.31.1 Aerobic and anaerobic expression of DadhA in *E. coli* BL21 (DE3) in LB medium supplemented with nitrate, molybdate, and ferric ammonium citrate

Aerobic overnight starter cultures were inoculated 1:100 into 50 mL of LB medium containing 1 mM sodium molybdate, 15 mM sodium nitrate, 1 mM ferric ammonium citrate, the appropriate antibiotic, and 0.5 mM dopamine. Cultures were grown at 37 °C. When the bacterial cultures reached OD$_{600}$, IPTG was added at a final concentration of 0.2 mM to induce expression. Cultures were incubated at 15 °C for 15 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.31.2 Aerobic and anaerobic expression of DadhA in *E. coli* BL21 (DE3) harboring the pPH149 plasmid in LB medium supplemented with nitrate, cysteine, molybdate, and ferric ammonium citrate

Aerobic overnight starter cultures were inoculated 1:100 into 50 mL of LB medium containing 1 mM sodium molybdate, 15 mM sodium nitrate, 1 mM ferric ammonium citrate, 1 mM L-cysteine, the appropriate antibiotic, and 0.5 mM dopamine. Cultures were grown at 37 °C. When cultures reached OD$_{600}$ = 0.600, IPTG was added at a final concentration of 0.2 mM to induce expression. Cultures were incubated at 15 °C for 15 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.31.3 Anaerobic expression of selected DadhA constructs in *E. coli* BL21 (DE3) adapting previous protocols for acetylene hydratase and arsenate reductase expression

*Conditions based on arsenate reductase expression:* Aerobic overnight starter cultures were inoculated 1:100 into 5 mL of LB medium containing 1 mM sodium molybdate. Cultures
were grown at 37 °C. Cultures were grown at 37 °C. When cultures reached OD₆₀₀ = 1, this culture was inoculated 1:100 into 50 mL of TB medium containing 10 mM magnesium sulfate, 1 % (w/v) glycerol, 2 mM ferric ammonium citrate, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM dopamine, the appropriate antibiotic, and 14 mM DMSO. IPTG was then added at a final concentration of 40 µM. Cultures were grown anaerobically at room temperature for 20 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

Conditions based on acetylene hydratase expression: Aerobic overnight starter cultures were inoculated 1:100 into 5 mL of LB medium containing 1 mM sodium molybdate. Cultures were grown at 37 °C. When cultures reached OD₆₀₀ = 1, this culture was inoculated 1:100 into 50 mL of TB medium containing 10 mM magnesium sulfate, 1 % (w/v) glycerol, the appropriate antibiotic, 2 mM ferric ammonium citrate, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM dopamine, and 50 mM fumarate. Cultures were grown at 37 °C. At OD₆₀₀ = 0.3, sodium nitrate was added (15 mM final concentration). Following one hour of incubation at room temperature, IPTG was added at a final concentration of 100 µM. Cultures were grown at 15 °C for 15 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.31.4 Multiple gene expression of Dadh in E. coli BL21 (DE3) ΔiscR and E. coli MG1655

Aerobic overnight starter cultures were inoculated 1:100 into 50 mL of TB medium containing 10 mM magnesium sulfate, 1 % (w/v) glycerol, 2 mM ferric ammonium citrate, the appropriate antibiotic, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM dopamine, and 14 mM DMSO. IPTG was then added at a final concentration of 40 µM. Cultures were grown anaerobically at room temperature for 20 hours. Cells were lysed using sonication, followed by
SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.3.5 Expression of DadhABC on pTrcHis2A in with and without auxiliary plasmids in E. coli

Aerobic overnight starter cultures were inoculated 1:100 into 50 mL of TB medium containing 10 mM magnesium sulfate, 1 % (w/v) glycerol, the appropriate antibiotic, 2 mM ferric ammonium citrate, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM dopamine, and 14 mM DMSO. IPTG was then added at a final concentration of 40 µM. Cultures were grown anaerobically at room temperature for 20 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.3.6 Expression of DadhABCE on pETDuet1 in E. coli MG1655(DE3)

Aerobic overnight starter cultures were inoculated 1:100 into 50 mL of TB medium containing 10 mM magnesium sulfate, 1 % (w/v) glycerol, the appropriate antibiotic, 2 mM ferric ammonium citrate, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM dopamine, and 14 mM DMSO. IPTG was then added at a final concentration of 40 µM. Cultures were grown anaerobically or aerobically at room temperature for 20 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed using LC-MS.

5.4.3.2 Purification of DadhA from E. coli BL21(DE3)

Aerobic overnight starter cultures of E. coli BL21 (harboring pET28a-dadhA_C-His6, pETDuet1_dadhE_N-His6, and pH149) were inoculated 1:100 into 2 different bottles of 2 L of aerobic LB medium containing the appropriate antibiotic, 1 mM sodium molybdate, 0.5 % (w/v)
glucose, 2 mM ferric ammonium citrate. Cultures were grown at 37 °C under shaking. At OD<sub>600</sub> = 0.500, L-cysteine (2 mM final concentration) and sodium nitrate (15 mM final concentration) were added. The cultures were then cooled to 15 °C for 1 hour, and then one of the bottles was sparged with nitrogen to generate an anaerobic environment. The bottle was tightly sealed with electrical tape following sparging. Finally, IPTG was added at a final concentration of 0.1 mM to both the aerobic and anaerobic cultures to induce expression, which took place at 15 °C for 15 hours. For purification of recombinant DadhA, the 2 L cultures were harvested by centrifugation, and the resulting pellet was resuspended in 30 mL 50 mM Tris pH containing 0.25M NaCl, followed by lysis using a cell disruptor (Avestin Emulsiflex C3). All of the clarified lysate was loaded onto 2 mL of HisPur Ni-NTA resin (Thermo Fisher Scientific, Waltham, MA) and eluted using a gradient of 50 mM to 200 mM imidazole (in 50 mM Tris pH 8 containing 0.25M NaCl). Fractions containing pure protein were combined and dialyzed over two rounds into in 50 mM Tris pH 8 containing 0.20 M NaCl and 10% w/v glycerol. The dialyzed protein was concentrated 10-fold using spin columns (VMR, catalog # 97027-9).

5.4.33 Expression of catechol dehydroxylase homologs in <i>E. coli</i> and <i>Shewanella sp. ANA-3</i>

Aerobic overnight starter cultures of <i>E. coli</i> or <i>Shewanella sp. ANA-3</i> were inoculated 1:100 into 50 mL of aerobic or anaerobic TB medium containing the appropriate antibiotic, 10 mM magnesium sulfate, 1 % (w/v) glycerol, 2mM ferric ammonium citrate, 1 mM L-cysteine, 1 mM sodium molybdate, 0.5 mM catechol substrate, and 14 mM DMSO. Arabinose was then added at a final concentration of 20 mM. Cultures were grown for 18 hours aerobically at 30 °C under shaking, or anaerobically at room temperature without shaking. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed by LC-MS.
5.4.34 Protein expression in *E. faecalis OG1RF*

To generate electrocompetent cells, aerobic overnight starter cultures in what medium were inoculated 1:100 into 100 mL BHI medium containing 0.5 M sucrose and 8% (w/v) glycine. These cultures were grown for 18 hours at 37 °C with no shaking, followed by harvesting by centrifugation at 4 °C. The cells were then resuspended in 10 mL of sterile, ice-cold deionized water containing 0.5 M sucrose and 10% (w/v) glycerol. Cells were pelleted by centrifugation and were resuspended in 1 mL of sterile, ice-cold deionized water containing 0.5 M sucrose and 10% (w/v) glycerol. 40 µL aliquots of these electrocompetent cells were then incubated with 4 µL of plasmid DNA for 5 minutes in a 2 mm gap electroporation cuvette. Cells were transformed in a MicroPulser electroporator (Bio-Rad) with a 2.5 kV pulse (time constant ~4.8 – 5.2), rescued with 0.9 mL of BHI medium containing 0.5 M sucrose and left on ice for 5 minutes. The cells were then incubated for 2 hours at 37 °C without shaking. Various dilutions were plated onto BHI medium supplemented with the appropriate antibiotic and incubated at 37 °C for 1-2 days.

For expression, anaerobic overnight starter cultures of *E. faecalis OG1RF* grown in BHI medium were inoculated 1:50 into 50 mL anaerobic BHI medium containing the appropriate antibiotic, 1 mM sodium formate, 0.25 mM dopamine, and 1 mM sodium molybdate. IPTG was then added at a final concentration of 0.1 mM. Cultures were grown for 18 hours aerobically at room temperature. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed by LC-MS.

5.4.35 Protein expression in *R. erythropolis* L-88

We adapted previously reported protocols for transformation and growth of *R. erythropolis* L-88 (97). All experiments were performed aerobically. To generate electrocompetent cells, we harvested 500 mL cultures grown at 28 °C, 150 rpm in in mid exponential phase (OD₆₀₀ ~ 0.4-0.6)
by centrifugation. We washed these cells once with 500 mL deionized water, followed by resuspension in 10 mL 30% PEG 1000 (solution in deionized water). These competent cells were used directly but could also be frozen at −80 °C in 400 µL aliquots for future use. For transformation, 50 ng of plasmid 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. 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 medium, and incubated for 4 hour at 28 °C, 175 rpm. Cells were plated onto LB agar plates+antibiotic (17 µg/mL chloramphenicol for pTipQC plasmid) and incubated at 28°C for 5–7 days. Single colonies were inoculated into 50–75 mL of LB +antibiotic (34 µg/mL chloramphenicol) and grown for 3–5 days at 28 °C, 175 rpm until reaching saturation. These were the starter cultures. For expression, the starter cultures were inoculated into 50 mL of LB medium (containing supplements and antibiotic) to a starting OD₆₀₀ of 0.2 and grown at 28 °C, 175 rpm. We used different types of LB medium. In one condition, the LB medium was supplemented with 0.25 mM ferric ammonium citrate, 0.5 mM L-cysteine, 0.5 mM sodium molybdate, 0.5 mM dopamine, and 4 mM sodium nitrate. In the other LB medium, the 4 mM sodium nitrate was substituted for 14 mM DMSO. When cultures reached an OD₆₀₀ of 0.6 (~6–8 hr), protein expression was induced with thiostrepton at a final concentration of 0.01 µg/mL, and cultures were incubated at 15 °C, 175 rpm for 20 hours. Cells were lysed using sonication, followed by SDS-PAGE analysis and setup of activity assays. Culture supernatants and lysate activity assays were analyzed by LC-MS.

5.4.36 Lysate assays to assess catechol dehydroxylase activity

Cultures were harvested by centrifugation and the pellet was resuspended in 1/50th of the original culture volume. Cells were lysed by sonication using the following protocol: 25%
amplitude, 1 min total sonication, 10 s on, 40 s off. 50 µL of lysate was mixed, in the following order, with 1 µL electron donors (final concentration 1 mM each of methyl viologen, 1 mM diquat dibromide, 1 mM benzyl viologen, all dissolved in water), 2 µL of sodium dithionite (2 mM final concentration, dissolved in water), and 1 µL of substrate (500 µM final concentration, dissolved in water). The assay mixtures were left at room temperature in an anaerobic chamber for 12–20 hours to allow dehydroxylation to proceed. Lysates were then diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS. Dadh constructs were incubated with dopamine and were analyzed by LC-MS using Method A described above. Hcdh and Dodh constructs were incubated with hydrocaffeic acid and DOPAC, respectively, and were analyzed by LC-MS using Method B described above. Tardh constructs were incubated with hydroxyhydroquinone and were analyzed by LC-MS using Method C described above.

5.4.37 LC-MS analysis of culture supernatants

Cultures were harvested by centrifugation and the supernatant was diluted 1:10 with LC-MS grade methanol and analyzed by LC-MS/MS. Dadh constructs were incubated with dopamine and were analyzed by LC-MS using Method A described above. Hcdh and Dodh constructs were incubated with hydrocaffeic acid and DOPAC, respectively, and were analyzed by LC-MS using Method B described above. Tardh constructs were incubated with hydroxyhydroquinone and were analyzed by LC-MS using Method C described above.
5.5 References


74. P. J. Ellis, T. Conrads, R. Hille, P. Kuhn, Crystal structure of the 100 kDa arsenite oxidase from Alcaligenes faecalis in two crystal forms at 1.64 A and 2.03 A. *Structure* **9**, 125-132 (2001).


Chapter 6. Concluding remarks and future directions
6.1 Concluding remarks

The human gut microbiota is one of the most densely colonized microbial habitats on Earth and is deeply intertwined with host biology (1). However, a major challenge facing gut microbiota research is the need to move beyond cataloging the organisms in this community to elucidating the mechanisms underlying their influence on human health (2). By altering the chemical structures of ingested and endogenous compounds, gut microbes can mediate the effects of host-derived molecules and xenobiotics (drugs, diet) on host physiology (2, 3). Interindividual variation in gut microbial metabolism poses major challenges in nutrition and pharmacology, and although gut microbial metabolic activities were identified as early as in the 1950s (4), few transformations have been connected to organisms, genes, and enzymes.

In this work, we have described the discovery and characterization of microbial strains, genes, and enzymes that metabolize neurotransmitters, drugs, and dietary compounds in the human gastrointestinal tract. We reported that the human gut microbe *E. faecalis* uses a promiscuous tyrosine decarboxylase (TyrDC) to degrade L-dopa, the main drug treatment for Parkinson’s disease. We also found that *tyrDC* predicts L-dopa metabolism by complex gut microbiota communities, potentially explaining the well-known interindividual variability in L-dopa efficacy (5-7). Additionally, we discovered that the co-prescribed human decarboxylase inhibitor carbidopa does not target microbial L-dopa decarboxylation, suggesting that some L-dopa metabolism may be unaccounted for in the current treatment regimen (8). The lack of carbidopa potency towards microbial metabolism inspired the development of the selective microbial decarboxylase inhibitor AFMT, which blocks L-dopa metabolism by *E. faecalis* and complex communities. AFMT could serve as a promising tool for understanding microbial contributions to drug metabolism in vivo. Finally, we discovered that the prevalent human gut bacterium *C. sporogenes* uses known enzymes
involved in aromatic amino acid metabolism (9) to transform L-dopa into m-hppa. This pathway is thought to make only minor contributions to L-dopa metabolism in vivo, but its clinical relevance is unclear (10). Discovery of genes and enzymes involved could enable further study of this alternative pathway for microbial L-dopa metabolism.

By studying the second step of the major pathway for gut microbial L-dopa metabolism (10), dehydroxylation of dopamine derived from L-dopa decarboxylation, we found that the prevalent gut Actinobacterium *E. lenta* dehydroxylates dopamine into *m*-tyramine using a catecholamine-specific, molybdenum-dependent dehydroxylase (Dadh). This discovery not only advanced our understanding of metabolism of Central Nervous System-targeted drugs in the human body, but also provided clues into the molecular basis of catechol dehydroxylation, a prominent transformation involved in the gut microbial breakdown of diverse host-derived molecules, drugs, and dietary compounds (10-13). Although catechol dehydroxylation ha been frequently reported in the scientific literature since its discovery in the 1950s, the genes and enzymes involved remained obscure (4, 13-21).

Combining microbiology, biochemistry, and bioinformatics, we uncovered that gut microbial catechol dehydroxylation is performed by a new class of molybdenum-dependent enzymes that are highly specialized in their transcriptional regulation and biochemical activity. These enzymes, a majority of which remain uncharacterized, likely enable gut microbial respiration on catechols as alternative electron acceptors. We also propose that catechol dehydroxylation proceeds by a unique reaction mechanism that differs from that of other characterized reductive aromatic dehydroxylases (22). Moreover, we found that catechol dehydroxylases are widespread among the microbiotas of other mammals and in the environment, where they participate in both primary and secondary microbial metabolism (23, 24). This suggests
that catechol dehydroxylation likely more prevalent than previously thought and underscores that the relevance of our work extends far beyond the human gut microbiota. Finally, we described progress towards further mechanistic and biochemical characterization of dehydroxylases. Consistent with our proposed mechanism for catechol dehydroxylation, we found that hydrocaffeic acid dehydroxylation reaction run in deuterated water incorporated high levels (>95%) of a single deuterium into the aromatic ring of the product m-hppa. Additionally, we attempted heterologous expression of catechol dehydroxylases from gut and environmental bacteria, laying the foundation for future biochemical and structural studies of a chemically and ecologically intriguing metalloenzyme family.

Our discovery efforts address a fundamental challenge facing gut microbiota research: assigning functions to the vast number of uncharacterized or misannotated genes present in human gut microbiomes. The gut microbiome encodes for over 3 million unique genes (25), yet 86% of the genes from HMP stool metagenomes cannot be assigned to known metabolic pathways, and half cannot be given even a homology-based annotation (26). Additionally, many members of large enzyme superfamilies also have unknown functions and are typically misannotated (27, 28). We have assigned candidate or biochemically confirmed functions to five previously uncharacterized gut microbial sequences, including four molybdenum-dependent enzymes. This has revealed a previously unappreciated functional diversity within bis-MGD enzymes (29), an enzyme family that plays diverse roles in environmental microbes but is best-known in the gut microbiota for its roles in Proteobacterial respiration and inflammation (30-32).

Bis-MGD enzymes may be far more important and in the human gut microbiota than previously thought, with a vast diversity of sequences remaining to be characterized. Surprisingly, the catechol dehydroxylases are largely confined to Actinobacteria such as *Gordonibacter* and
*Eggerthella* in the human gut microbiota. Our analysis suggests that the previously reported unusually high number of bis-MGD enzymes within these microbes is partially explained by an expansion of the catechol dehydroxylase class (33, 34). This raises the possibility that distinct bis-MGD enzyme family members may have evolved within different microbial clades and raises questions about whether other enzyme classes have similarly specialized roles. In addition to enhancing our fundamental understanding of the gut microbial metabolic repertoire, enzyme discovery efforts such as ours enable more accurate annotation of microbiome sequence data and provide new insight into the evolution and distribution of enzyme superfamilies (27, 28, 35, 36).

Our studies also reinforce the emerging idea that that gut microbial phylogeny is often not predictive of microbial function (37-40). For example, we found that complex gut microbiota communities and closely related human gut Actinobacteria and display remarkable variability in the removal of a single hydroxyl group from different catechols. This was generally explained by variable gene presence, highlighting the need to move beyond phylogeny to understand microbiota function. However, in the case of the dopamine dehydroxylase, a single nucleotide polymorphism, rather than gene presence, predicted activity by strains and complex gut microbiota communities. Even non-metabolizing strains express the dopamine dehydroxylase in response to dopamine, suggesting that activity is controlled at the protein level. This highlights a frequently overlooked possibility in microbiome research: that simply detecting genes and their expression may not sufficiently explain the activities encoded by the human gut microbiome. To meaningfully predict metabolism from ever-increasing amounts of microbiome sequencing data, we may need to go beyond genes and enzymes and understand the full spectrum of genetic, biochemical, and ecological factors contributing to a metabolic activity.
Gut microbial metabolism is a two-way interaction between the microbiota and the host. Many studies seek to understand study microbial metabolism in light of potential host effects (13, 39, 41-43), but for a majority of known microbial transformations, including those where enzymes have been reported (11, 37), the physiological role of the activity for the metabolizing microorganism still remains obscure. Our studies have provided rare insight into how microbial metabolism of drugs and non-polysaccharide dietary compounds could potentially affect microbial community composition. Informed by our discovery of TyrDC and its role in L-dopa metabolism, we found that L-dopa decarboxylation promotes acid survival of *E. faecalis*. Similarly, we found that dopamine promotes *E. lenta* growth by serving as a terminal electron acceptor. Our understanding of the dopamine dehydroxylase Dadh allowed us to directly tie this growth benefit to active dopamine dehydroxylation in isolated cultures, competition experiments, and minimal communities. Finally, by targeting Dadh using PCR, we demonstrated that this growth benefit extends to native, complex human gut microbiota communities, linking the transformation to a clear phenotype in an ecologically relevant context. These data highlight the importance of enzyme discovery for understanding how metabolism impacts gut microbes. Ultimately, characterizing microbial genes and enzymes not only helps dissect the physiological role of prominent transformations, but could also explain known links between dietary compounds, pharmaceutical drugs, and neurotransmitters, and microbial community structure and microbial growth (44-46).

Our work has implications for medicine. Understanding the molecular basis of microbial metabolism uncovers potential biomarkers that could inform drug development, clinical trial design, and clinical practice. Knowledge of gut microbial metabolism could shape clinical trial design and clinical practice by allowing physicians to screen for detrimental or beneficial activities before prescribing drugs (Figure 6.1). For example, *tyrDC* or *E. faecalis* abundances, which predict
L-dopa metabolism by complex gut microbiota communities, could be considered as a factor in determining L-dopa dosage in Parkinson’s disease patients. Additionally, understanding which functional groups are prone to microbial metabolism could allow medicinal chemists to either avoid these structural features or incorporate them into prodrugs to enable selective drug activation in the GI tract (47). For example, the catechol functional group, which is extensively metabolized by the gut microbiota, could be avoided in drug design, or rationally incorporated in a pro-drug that would require microbial activation into a pharmacologically active dehydroxylated metabolite. However, further work in human patients is necessary to establish the distribution of the genes that we have discovered and their associated metabolic activities, and how these relate to host physiology. Long-term, understanding the distribution of specific microbial metabolic functions within human populations could significantly advance personalized medicine and microbiota-targeted drug design.

A molecular understanding of microbial metabolism also has implications for nutrition (Figure 6.1). Diet is a cornerstone of human health. Though many epidemiological studies linking dietary patterns with health outcomes have yielded conflicting results, few have taken gut microbial metabolic capabilities into account. For example, while consumption of polyphenols from fruits and vegetables is generally associated with health, effects can be variable between people and between studies (14, 48). This could potentially be explained by differences in microbial metabolism, but a lack of molecular information has made it difficult to link the gut microbiota to potential health effects. We have identified genes involved in the metabolism of commonly consumed dietary phytochemicals. We anticipate that these microbial genes could be incorporated into clinical studies to explore potential relationships with host nutrition and diet-induced shifts in microbial community composition (49). Ultimately, a comprehensive molecular
understanding of how gut microbes process dietary components could enable rational design of “functional foods” or prebiotics to treat conditions such as metabolic disease and malnutrition. Microbial metabolism could also inform personalized nutrition, in which diets are individually customized for patients’ metabolic profiles and gut microbiotas (Figure 6.1) (50).

Figure 6.1 Implications of understanding gut microbial metabolism of xenobiotics. Incorporating a mechanistic understanding of microbial transformations, along with knowledge of host genetics and metabolism, could A) enable personalized medicine and B) inform personalized nutrition.
Our studies also provide rare molecular insight into the emerging connection between the gut microbiota and the host nervous system. Microbiota features are associated with mental health disorders in human patients (51, 52), and studies in germ-free, conventional, and gnotobiotic animals indicate a potentially causal role for gut microbes in nervous system development and function (53-57). However, the mechanisms underlying microbiota-nervous system interactions are poorly defined (58). Our characterization of Dadh uncovered a previously unappreciated role for gut microbes in catecholamine metabolism, revealing that the prevalent Actinobacterium *E. lenta* specifically senses and metabolizes abundant gut neurotransmitters (59, 60), potentially for use in anaerobic respiration. Along with recent work on serotonin and GABA (46, 52), our discovery suggests that gut microbes have evolved to specifically interact with the host nervous system. This critical molecular information, which has been conspicuously absent from many previous microbiota studies of the gut-brain axis (55, 57, 58), sets stage for mechanistic, hypothesis-driven investigations of microbiota-neurotransmitter interactions and their relevance for host physiology. For example, one could explore interactions between Dadh and catecholamine-related phenotypes such as gut motility, brain function and microbial pathogenesis (51, 55, 61, 62). Uncovering the mechanisms by gut microbes recognize and catabolize neurotransmitters could provide insight into host nervous system development and function, potentially serving as biomarkers for human health.

In addition to revealing microbial genes relevant to human health, our work highlights that functionally characterizing microbial genes and enzymes can reveal novel therapeutic targets (Figure 6.1). Parkinson’s disease is a debilitating neurological condition affecting millions of people worldwide (63). L-dopa therapy significantly improves symptoms but suffers from high interindividual variability in both efficacy and side effects (7, 64-66). Given the emerging role of
the gut microbiota in L-dopa efficacy established by us and others (51, 67, 68), gut microbial metabolism of this drug could represent a novel therapeutic target (Figure 6.1). Through biochemical characterization of both bacterial and human decarboxylases, we identified a small-molecule inhibitor (AFMT) that inhibits microbial L-dopa metabolism in minimal and complex communities and increases serum L-dopa in gnotobiotic mice. While further work is needed to determine the in vivo effects and clinical promise of this molecule, microbiota-targeted small molecules more generally represent an exciting mode of manipulating microbial metabolism (69-74). This strategy overcomes the limitations of genetic manipulation, having the potential to target a single metabolic activity across diverse species. Molecules like AFMT will serve as useful tools for elucidating effects of microbial metabolism on both host and the microbiota. Long-term, such molecules could inspire development of novel microbiome-targeted therapeutics.

The relevance of our work extends far beyond the human body. We found that the chemical strategies used to enable microbial survival and interactions in the human gut are relevant to a broad range of species and habitats. While mammalian gut microbiomes have previously been compared in terms of gene content and species composition (75, 76), our work on catechol dehydroxylases provides functional evidence for conservation of specific gut microbial metabolic pathways across distinct hosts. While this hints at potentially important roles for catechol dehydroxylation across mammalian gut communities, the distribution of putative dehydroxylases among environmental microbes suggests this chemistry is present in many additional microbial habitats. We also found that the L-dopa decarboxylating enzyme TyrDC is widely distributed among Enterococci from diverse environments, further highlighting the wide distribution of enzymes that are initially discovered or studied in the human gut microbiota. These data reinforce findings from studies of additional gut microbial enzymes. For example, gut microbial
carbohydrate-degrading enzymes and glycyl radical enzymes, which play important roles in degrading diet-derived polysaccharides, amino acids, and osmolytes in the human gut, are also found in environmental isolates (27, 35, 38, 77-79). Thus, enzyme discovery in the human gut microbiota not only has implications for improving human health and disease, but also for discovering novel catalytic functions and metabolic pathways broadly relevant to microbial life. Because the gut microbiota is an unexplored microbial habitat with a large portion of “genetic dark matter”, we anticipate that studies similar to ours will continue to uncover novel chemistry and enzyme classes (37, 39).

Finally, our findings provide a framework for linking metabolic transformations performed by complex gut microbial communities to individual strains, genes, and enzymes. We have demonstrated the utility of combining both modern and traditional approaches to dissect microbial function. For example, our combination of modern transcriptomics and genomics tools with classical activity-based purification enabled the discovery and characterization of Dadh, the founding member of the catechol dehydroxylase enzyme class. The approaches used to discover Dadh may prove particularly useful in future discovery and characterization efforts involving genetically intractable bacteria. Finally, our broad exploration of a class of metabolic transformations contrasts with the more common focus on metabolism of individual drugs or dietary compounds (27, 35, 37, 38, 43, 67, 78, 80, 81). This functional group-focused approach may greatly increase the efficiency with which we can link metabolic activities to microbial genes and enzymes. We envision that related experimental workflows could find broad utility in the discovery of gut microbial enzymes catalyzing other widespread, biologically significant reactions, including reductive metabolism of additional functional groups that are prevalent in diverse molecules encountered by microbes in the human gut (82). Overall, our study showcases
the value of incorporating enzyme discovery and characterization efforts into investigations of gut microbial metabolism, setting the stage for investigations of the chemical mechanisms and biological consequences of prominent microbial transformations in the human body and beyond.

6.2 Future directions

Below we highlight specific recommendations for immediate and long-term future directions building upon our current work.

6.2.1 Evaluate microbial contributions to L-dopa efficacy in vivo

We have described organisms and enzymes that degrade the Parkinson’s drug L-dopa, but we have not yet established the precise contribution of microbial metabolism to drug efficacy and side effects in the host. To establish the clinical relevance of microbial L-dopa metabolism, we propose further studies in animal models and patient cohorts. First, gnotobiotic mice could be colonized with wild-type E. faecalis or an isogenic tyrDC mutant. These animals would receive L-dopa with or without carbidopa, followed by pharmacokinetic profiling of drug levels in serum and host tissues. This critical experiment would test whether the TyrDC enzyme impacts circulating L-dopa levels. Given that our preliminary pharmacokinetic experiments (assessing AFMT effects) involved the ratio of 4:1 carbidopa to L-dopa, future studies should explore other dose regimens, including the clinically relevant ratio of 1:4 carbidopa to L-dopa (83), as well as the method of L-dopa delivery, including increasingly popular slow-release formulations (84). It would also be important to consider the animal host; whereas we used mice, all previous studies of microbial L-dopa metabolism used rats (85, 86). Drug pharmacokinetic profiles can differ between rats and mice (87, 88). Although such species differences have not been explored in L-dopa metabolism, it is worth considering for future studies. Further gnotobiotic work could also include defined, multi-species communities, which more closely reflect the native ecological
context (89, 90). Such communities may overcome the lack of dopamine dehydroxylation in our preliminary metabolic cage experiments with mice co-colonized with both *E. faecalis* and *E. lenta.* To link microbial metabolism to changes in drug efficacy, experiments could be extended into animal models of Parkinson’s disease, although not all these models are L-dopa responsive (55, 91, 92).

It will also be important to conduct patient studies, including correlational analysis in existing cohorts. Whereas diverse studies have explored relationships between gut microbiota composition and Parkinson’s disease (93-99), we are unaware of studies directly linking microbiome profiles to L-dopa efficacy and side effects. Researchers can now search for relationships between *E. faecalis, E. lenta, tyrDC,* and *dadh* abundance, and variables related to L-dopa efficacy, including on/off times, drug dosage, and common side effects (68, 100). In addition to stool gene abundance, microbial L-dopa metabolism ex vivo should be considered as a correlate of patient response. Given that antibiotics have been reported to improve L-dopa efficacy and elevate drug serum levels in patients (68, 101), studies could move beyond correlation to perform controlled intervention studies. For example, a study could include Parkinson’s patients dosed with antibiotics or a placebo control, followed by immediate assessment of L-dopa efficacy and pharmacokinetics. Microbiome composition and metagenomes data could be collected in parallel to explore any predictors of the response to antibiotic intervention. Taken together, these proposed clinical studies would provide missing critical information about microbial L-dopa metabolism and its effect on drug pharmacokinetics and efficacy in the host.

**6.2.2 Further characterize AFMT and develop novel analogs**

Our study warrants further characterization of AFMT. While we found that AFMT did not inhibit growth of *E. faecalis* and *E. lenta,* its impact on the broader gut microbial community are
less clear. Further work could assess its impact on gut microbiota composition in conventional animals or complex gut microbiota cultures ex vivo. Given that TyrDC is important for *E. faecalis* survival under acid stress, AFMT could have unanticipated effects in vivo, particularly in low pH compartments of the gastrointestinal tract. It will also be important to assess whether AFMT inhibits L-dopa decarboxylation by other decarboxylating organisms beyond *E. faecalis*, including *E. faecium*. Finally, although AFMT is a promising tool compound, it suffers from the potential drawback that it is sensitive to host metabolism. Specifically, AFMT can be hydroxylated at the *meta*-position of the aromatic ring by the host enzyme tyrosine hydroxylase, generating α-fluoromethyl-L-dopa (AFMD), a potent inhibitor of the host L-dopa decarboxylase AADC (102-105). Because L-dopa crosses the blood-brain barrier, AFMD is expected to inhibit L-dopa decarboxylation in both peripheral tissues and in the brain, causing toxicity and side effects. Additionally, host metabolism of AFMT would complicate the interpretation of the potential inhibitory effects of AFMT in vivo, as apparent decreases in peripheral decarboxylation could result from AFMD. To overcome potential off-target effects of AFMT, further efforts should focus on developing AFMT analogs, changing the aromatic ring substituents to remove potential susceptibility to host metabolism, and evaluating with different PLP-targeting warheads (106). Solving crystal structures of AFMT-bound *E. faecalis* TyrDC and gaining a broader understanding of TyrDC’s substrate scope would greatly aid in future inhibitor development.

### 6.2.3 Understand how catechol dehydroxylation impacts host and gut microbiota in vivo

Many questions remain about the physiological impacts of catechol dehydroxylation on both host and gut microbiota. Our studies of Dadh revealed a striking biochemical and transcriptional specificity for catecholamines. Dopamine and norepinephrine are both produced by the host in the gastrointestinal tract and are available to gut bacteria (59, 60), but the precise
mechanisms by which gut bacteria sense and import these neurotransmitters is unclear. Additionally, the function and distribution of catecholamines in the gastrointestinal tract has received far less attention than the neurotransmitter serotonin (46, 53). Thus, our work highlights a need for further study of the gastrointestinal catecholaminergic system. Newly developed tools, such as single-cell sequencing and cell labeling using adeno-associated viruses (AAVs) should enable physical mapping of neuronal and non-neuronal processes and illuminate potential interactions with gut microbes (107, 108).

Additionally, it will be important to investigate links between catechol dehydroxylase genes, their associated metabolites, and host health in human patients. Coupled with investigations of the potential host targets of dehydroxylated metabolites in cell cultures or receptor-based screens (109, 110), such investigations could establish the extent to which catechol dehydroxylation directly impacts host physiology. Finally, while we have demonstrated that catechol dehydroxylation confers a growth advantage in complex communities ex vivo, the relevance of this metabolism to microbiota composition in vivo is still not clear. To address this gap, our competition and minimal community growth experiments could be replicated in gnotobiotic animals dosed with purified compounds or intact foods (42). Given that many catechol dehydroxylases metabolize dietary compounds, one could also track the abundance of catechol dehydroxylating organisms and enzymes in response to dietary interventions in humans (49, 111).

6.2.4 Understand the molecular basis of catechol dehydroxylase regulation in human gut microbes

Our data support a model in which gut microbial catechol dehydroxylase expression is tightly regulated by individual enzymes’ substrates. The molecular basis of this specificity in transcriptional regulation represents an intriguing avenue for future research. Genes encoding for
catechol dehydroxylases frequently co-localize with putative transcriptional regulators. These regulators may sense catechol substrates and specifically regulate dehydroxylase expression. To confirm the roles of the regulators, future efforts could follow common approaches used in biosensor characterization (112), cloning the regulators and the catechol dehydroxylase promoter sequence for expression in a heterologous host. The dehydroxylase promoter could drive expression of a fluorescent protein, potentially generating a signal in the presence of the catechol substrate. Uncovering the molecular basis of catechol dehydroxylase regulation would enhance our fundamental understanding of a widespread and clinically relevant microbial activity, inform new approaches to identify enzyme-substrate pairs, and ultimately guide the design of biosensors for endogenous and xenobiotic compounds.

**6.2.5 Characterize catechol dehydroxylase diversity in the human gut and beyond**

We initially became interested in catechol dehydroxylation because of its prominence in the human gut microbiota. By discovering and biochemically characterizing a single, gut bacterial catechol dehydroxylase, we eventually uncovered that catechol dehydroxylation is present in other mammalian microbiotas and diverse environmental bacteria. Intriguingly, closely related sequences are shared across microbial phylogeny, habitat, and lifestyle. For instance, catechol dehydroxylases in a drug-metabolizing human gut Actinobacterium (37) are related to those found in a hyperthermophilic sulfate-reducing Archeon (113) and a soil bacterium well-known for its ability to dehalogenate aromatic compounds (114). To me, the author, these are perhaps the most fascinating findings of this thesis. The vast majority of this catechol dehydroxylase diversity remains uncharacterized. To uncover the functions of the many uncharacterized dehydroxylases, future work could utilize high-throughput culture-based screens and transcriptomics, similar to the strategies that enabled identification of the dopamine dehydroxylase and its gut microbial
homologs. Additionally, development of genetic tools for intractable gut Actinobacteria would enable more rapid pairing of substrate-enzyme pairs. In environmental bacteria where genetic tools are available, catechol dehydroxylase genes could be knocked out or expressed in the native host, enabling assessment of biochemical activity and physiological effects (23, 24). These studies would answer current questions about the evolution and function of a new enzyme family and highlight its potential ecological roles.

Further discovery efforts of uncharacterized dehydroxylases should consider molecules beyond commercially available catechols as potential substrates. For example, in Chapter 4, we suggested that some substrates may first encounter the microbiota in the form of methoxylated aromatics, which could be demethylated by the microbial community and liberate catechols subject to dehydroxylation (Figure 6.2). This type of metabolism is known in the gut microbial metabolism of the dietary lignan pinoresinol (11). However, future efforts should consider non-catecholic substrates that could undergo the arrow pushing that we propose is necessary for dehydroxylation. Such substrates include aromatic compounds carrying a tri-hydroxyl substitution pattern (pyrogallol), which were accepted by Dadh in vitro (see Chapter 4 for details) and also undergo dehydroxylation in the metabolism of the FDA-approved drug Fostamatinib and diet-derived ellagic acid (13, 15). The presence of trihydroxylated aromatics in many plants indicates these molecules have the potential to interact with microbes (115). Another possibility is 1,4-dihydroxylated aromatic compounds (hydroquinones), which could have a similar ability as catechols to undergo the tautomerization and re-aromatization that we propose is involved in catechol dehydroxylation (Figure 6.2). Consistent with this proposal, hydroquinone dehydroxylation was previously observed by Desulfitobacterium hafniense, an environmental organism that harbors many catechol dehydroxylase homologs, and by complex microbial
communities from freshwater sediments and sewage sludge (116, 117). Other substrates include hydroxylated indole compounds, such as serotonin and indican, which should have the ability to be dehydroxylated by our proposed arrow pushing mechanism for catechol dehydroxylation (Figure 6.2). Finally, substrates could include non-aromatic hydroxyketones (Figure 6.2). Hydroxyketones feature prominently in anaerobic microbial metabolism of aromatic amino acids, but we are unaware of examples of the proposed direct dehydroxylation of α-hydroxyketones to produce the corresponding ketone. In summary, we propose that further discovery efforts involving members of the catechol dehydroxylase enzyme class should consider all of these non-catecholic molecules as potential substrates.
Figure 6.2 Potential substrates for uncharacterized dehydroxylases.
This figure displays the general reaction pathway for potential substrates to undergo dehydroxylation, as well as specific examples of potential molecules. Examples harboring red hydroxyl groups are known to be dehydroxylated at that position. A urolithin M6 and a derivative of pinoresinol are both dehydroxylated by *Gordonibacter* from the human gut microbiota. TCHQ is dehydroxylated by *D. hafniese*, an environmental organism that harbors many uncharacterized dehydroxylase homologs.
6.2.6 Characterize the mechanism of catechol dehydroxylation

Ultimately, many of the directions highlighted above will be enhanced by a greater understanding of the fundamental chemistry involved in catechol dehydroxylation. The substrate scope of Dadh and the lack of sequence homology to the only other known reductive aromatic dehydroxylase, 4-HBCR, suggests that catechol dehydroxylation proceeds by a unique mechanism. This proposed mechanism was consistent with the high (>95%) deuterium incorporation into dehydroxylated metabolite \(m\)-hppa when hydrocaffeic acid dehydroxylation was performed in deuterated water. However, further work is needed to understand the details of this chemically challenging and intriguing reaction. Future efforts could incorporate spectroscopy, kinetics, and reactions with substrate analogs to investigate the mechanistic proposal presented in this work. These mechanistic studies would be greatly enhanced by structural characterization of a catechol dehydroxylase. A full protein structure would reveal the nature of the molybdenum cofactor and its ligands, reveal residues involved in catalysis and active site access, and enable homology-based prediction of additional dehydroxylase structures. Importantly, however, further mechanistic work depends on progress in heterologous expression or optimization of native purification protocols to increase protein yield and quantity.
6.3 References


