Heme Homeostasis and Utilization in *Clostridioides difficle* Pathogenesis and Persistence

By

Reece Jeffrey Knippel

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

January 31, 2020

Nashville, Tennessee

Approved:

David M. Aronoff, M.D.

Maria Hadjifrangiskou, Ph.D.

Richard M. Peek, Jr, M.D.

Steven D. Townsend, Ph.D.

Eric P. Skaar, Ph.D., M.P.H.

ACKNOWLEDGMENTS

To my mentors, colleagues, and collaborators who provided the guidance, expertise, and support at Vanderbilt, this work would not have been possible without you. First and foremost, I would like to thank my dissertation mentor, Dr. Eric Skaar. He took a chance on a chemist from Idaho and with his guidance I have grown into the scientist I am now. He also took a risk by allowing me to become the second person in his lab to work on *C. difficile* where, under his watchful eye, I was able to develop the following dissertation from the ground up. This experience has allowed me to build confidence in myself and gain pride my work. Thank you for trusting me, being my scientific role model, and supporting me not only in the day to day but allowing me to showcase my work at conferences around the nation.

I would like to thank the Skaar lab family for providing answers to my questions, support in the face of any obstacle, and critical yet constructive feedback on my work. I have to specifically thank members of "team C. diff" including the original member Joe Zackular, my across the bay mate Chris Lopez, Maribeth Nicholson, and Aaron Wexler for serving as constant fountains of advice and soundboards for my work. I'd like to thank all my current and former science siblings that I had the pleasure to overlap with my time in the lab. A special thanks to Lauren Palmer, Zach Lonergan, and Clare Laut for providing years of friendship and comradery. I thank Andy Wiess for his help with the RNA-seq analysis in chapter II. I would also like to thank Nichole Lobdell and Nichole Maloney as well as all of the research assistants that are able to keep everything in the lab functioning and providing support for my work.

To my thesis committee members: Thank you for your insight and expansion of my scientific endeavors over the past few years. Thank you to Dr. Dave Aronoff for serving as my committee chair and always providing enthusiastic support while also critical and direct feedback. Also, thank you for providing the anaerobic chamber where the majority of the *in vitro* work was performed in this dissertation. Thank you Dr. Rick Peek for always reminding me to consider the application of my work to patients and providing the pilot funding for the *C. difficile* work. Thank you Dr. Maria Hadjifrangiskou for providing the insights of the intricate aspects of microbiology. Finally thank you to Steve Townsend for providing the perspective of a scientist not entrenched in the microbial dogma. You all have been instrumental in my graduate education.

I would like to thank all of my collaborators and cores who have helped me immensely on this work. I'd like to thank Drs. Jess Moore and Richard Caprioli for the imaging mass spectrometry. Thank you to Arianna Celis and Dr. Jennifer DuBois for the intracellular heme analysis. I am grateful to Dr. Kay Washington and the Tissue Pathology Shared Resource Core for blinded scoring and preparation of the pathology slides. Thank you to Dr. Valérie de Crécy-Lagard for the taxonomic analysis of HsmA. I also thank Dr. Seth Walk for providing *C. difficile* strain R20291 and Dr. Joe Sorg for providing the *Bacillus subtilis* mating strain, *C. difficile* plasmids, and guidance for the creation of the *C. difficile* mutant strains.

I am grateful for the financial support towards my research training and the career opportunities provided at Vanderbilt. Thank you to the Chemistry-Biology Interface training grant and Dr. Gary Sulikowski for providing training and support during my second and third year of graduate school. I would also like to thank the Molecular Microbial

Pathogenesis training grant and Dr. Fernando Villalta for providing support during my last few years. I am also thankful to the grants and funding awarded to Dr. Eric Skaar that directly funded this working including: the National Institute of Allergy and Infectious Diseases grant R01AI073843, the National Institute of Diabetes and Digestive and Kidney Diseases grant P30DK058404, and the Ernest W. Goodpasture Chair in Pathology. I appreciate my director of graduate studies Dr. Chris Aiken and Lorie Franklin for their support. Thank you to the BRET office and Dr. Ashley Brady for the quality career advice and opportunities to train as a professional and identify scientific careers.

Finally, I would like to thank my friends and family who have and are currently supporting me through this seemingly endless pursuit of education. Thank you to my mother and father who despite everything have always supported my career goals. Immense gratitude to my friends who have not only been there in times of triumph but also for the support in times of trail. And to Nicole Putnam, my soon to be wife, thank you for providing endless support during this process, always keeping me in check, being my perfect counterpart, and being the person who makes me strive to be the absolute best version of myself I can be. We survived graduate school together and I know we can survive anything else. My time in Nashville has been transformative and I thank and love you all for helping me become who I am today.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS x

CHAPTER

I. INTRODUCTION 1
Clostridioides difficile 1
C. difficile induced gastrointestinal inflammation 2
The heme paradox: utilization and toxicity 4
C. difficile and heme 7
Conclusions 8

II. HEME SENSING AND DETOXIFICATION BY HATRT CONTRIBUTES TO PATHOGENESIS DURING *CLOSTRIOIDES DIFFICILE* INFECTION 10
Introduction 10
Materials and methods 10
Ethics statement 10
Bacterial strains, growth conditions, and plasmids 11
RNA extraction and sequencing 16
Quantitative RT-PCR 18
Polyclonal antibody generation 18
Immunoblotting analysis 18
Protein expression and purification 19
Absorption spectroscopy 20
XylE reporter assays 20
LC-MS heme quantification 21
Mouse model of CDI 23
C. difficile toxin cytotoxicity determination 24
Statisical Analysis 25
Results 25
C. difficile infection increases hemoglobin abundance in the gastrointestinal lumen 25
The transcriptional response of *C. difficile* to heme exposure 29
The *hatRT* operon increases expression in response to heme and confers heme resistance 32
HatR functions as a transcriptional repressor of the *hatRT* operon 37
HatT reduces intracellular heme concentrations 42

HatT promotes pathogenicity in a mouse model of *C. difficile* infection . 44
Discussion 46

III. *CLOSTRIDIOIDES DIFFICILE* HIJACKS HOST HEME FOR INCORPARTION INTO AN OXIDATIVE STRESS DEFENSE SYSTEM 50
Introduction 50
Materials and methods 51
Bacterial strains and plasmids 51
hsmR::CT and *hsmA::CT* strain generation 53
Complementation plasmids 53
Protein expression plasmids 54
Heme, paraquat, and antibiotic toxicity *C. difficile* growth assays 54
qRT-PCR 56
HsmR multiple sequence alignment 57
HsmA taxonomic distribution and physical clustering 57
Protein expression and purification 58
Solubilization of membrane fractions 59
Absorption spectroscopy 59
Heme and oxidative stress toxicity *S. aureus* growth assays 60
RNA-sequencing analysis 60
Relapse mouse model of CDI 61
Bacterial burden determination 61
Reactive oxygen species measurements 62
Results 62
Transcription of the *hsmRA* operon occurs rapidly in response to heme .. 62
HsmR and HsmA reduce heme toxicity 65
HsmR and HsmA coordinate to reduce heme toxicity through sequestration 67
HsmR is a transcriptional activator of the *hsmRA* operon 73
HsmA employs exogenous heme to confer resistance to oxidative stress and antibiotics 76
HsmR and HsmA promote resistance to vancomycin during CDI 81
Discussion 88

IV. SUMMARY AND FUTURE DIRECTIONS 93
Heme sensing, utilization, and detoxification in *C. difficile* 96
Defining anaerobic heme toxicity in *C. difficile* 97
Structural and mechanistic investigation of HatRT and HsmRA 98
Further elucidation of the contribution of HatRT and HsmRA to *C. difficile* pathogenicity 99

REFERENCES 103

APPENDIX
A. Tables associated with chapter II 119
B. Tables associated with chapter III 133

LIST OF TABLES

Table Page

1. Bacterial strains and plasmids used in chapter II..... 14

2. Oligonucleotides used in chapter II..... 15

3. Bacterial strains and plasmids used in chapter III..... 52

4. Oligonucleotides used in chapter III 55

LIST OF FIGURES

Figure Page

1. *C. difficile* induced inflammation of the gastrointestinal epithelium 3

2. Bacterial mechanisms to reduce heme toxicity 6

3. Hemoglobin accumulates in the cecum during *C. difficile* infection 27

4. Additional images of hemoglobin accumulation in the cecum during C. difficile infection 28

5. *C. difficile* encodes machinery to detoxify heme 30

6. The *hatRT* operon responds to and relieves heme toxicity 34

7. Complementation of *hatR::CT* and *hatT::CT* heme sensitivity 36

8. HatR transcriptional repression of the *hatRT* operon is released through direct heme binding 39

9. Purified recombinant HatR and HatR H99L 40

10. Histidine residues within HatR that are not required for heme binding 41

11. HatT reduces intracellular heme concentrations 43

12. *hatT::CT* displays reduced pathogenicity in a mouse model of CDI 45

13. *hsmRA* transcriptionally responds to and detoxifies heme 64

14. Heme sensitivity in WT, *hsmR::CT, hsmA::CT*, and complemented strains 66

15. HsmR binds and senses heme and HsmA reduces heme toxicity through sequestration 69

16. Purified recombinant HsmR and HsmR H50A 71

17. Conserved histidine residues within HsmR that are not required for heme binding 72

18. HsmR acts as an activator of the *hsmRA* operon 74

19. The regulon of HsmR is limited to the *hsmRA* operon 75

20. *hsmR::CT* and *hsmA::CT* are sensitive to paraquat in the presence of heme 78

21. HsmA reduces oxidative stress 79

22. HsmA is not involved in protection against hydrogen peroxide..... 80

23. The *hsmRA* operon decreases sensitivity to vancomycin in the presence of heme during infection..... 83

24. Protoporphyrin-IX does not rescue vancomycin toxicity..... 84

25. HsmRA decreases sensitivity to metronidazole in the presence of heme 85

26. HsmA is conserved among multiple bacterial species 87

27. Physical clustering of genes encoding HsmA and the heme efflux proteins HtrAB..... 91

28. Heme sensing, utilization, and detoxification in *C. difficile*..... 96

LIST OF ABBREVIATIONS

BHIS	Brain-heart-infusion broth supplemented with yeast extract
CDI	*Clostridioides difficile* infection
CDMM	*C. difficile* minimal media
CFU	Colony forming units
DHR-123	Dihydrorhodamine 123
H&E	Hematoxylin and eosin
HatR	Heme activated transporter regulator
HatT	Heme activated transporter
HrtAB	Heme-regulated transporter
HO	Heme oxygenase
HssRS	Heme sensing system regulator and sensor
IPTG	Isopropyl-1-thiol-D-galactopyranoside
Isd	Iron-regulated surface determinant
Kat	Catalase
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
MALDI IMS	Matrix-assisted laser desorption ionization imaging mass spectrometry
OD	Optical density
PBS	Phosphate buffer saline
PPIX	Protoporphyrin IX

qRT-PCR	Quantitative reverse transcription polymerase chain reaction
R20291	*C. difficile* R20291
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RPKM	Reads per kilobase per million mapped reads
rRNA	Ribosomal RNA
RNA-seq	RNA sequencing
SOD	Superoxide dismutase
TBS	Tris-buffered saline
TCCFA	Taurocholate cycloserine cefoxitin fructose agar
TcdA	*Clostridioides difficile* toxin A
TcdB	*Clostridioides difficile* toxin B
TFA	Trifluoracetic acid
TSA	Tryptic soy agar
TSB	Tryptic soy broth
WT	Wild-type

CHAPTER I

INTRODUCTION

Clostridioides difficile

Clostridioides difficile (formerly *Clostridium difficile*) is a Gram-positive, spore forming obligate anaerobe that is an urgent threat to public health as it is the leading cause of nosocomial diarrhea in the United States with half a million infections resulting in 29,000 deaths per year [1]. In healthy individuals, the gastrointestinal microbiome provides colonization resistance against *C. difficile* infection (CDI) [2]. Ingestion of *C. difficile* spores following perturbation of the microbiome, commonly due to antibiotic therapy in a healthcare setting, results in germination into vegetative cells in the small intestine after sensing primary bile acids, such as taurocholate, and the amino acid glycine [2]. Vegetative *C. difficile* colonizes the colon, causing a wide range of diseases that vary from infectious diarrhea to pseudomembranous colitis [3]. However, in some cases the pathogen is cleared or asymptomatically carried [3]. The symptoms of disease result from the production of two potent toxins, toxin A (TcdA) and toxin B (TcdB), which function as glycosyltransferases that inactivate specific Ras and Rho GTPases leading to the disruption of host cell function [4]. Throughout infection, *C. difficile* continues the production of spores perpetuating the spread of disease [2]. The initial treatment for CDI consists of an antibiotic regimen of vancomycin, metronidazole, or fidoxamicin [5]. For reasons that remain unknown, 1 in 5 patients will suffer from recurrence requiring further antibiotic treatment [5]. Patients with serious cases of multiple recurrence may benefit from a fecal microbiota transfer [6]. Through either treatment option, the gastrointestinal microbiome recovers and restores

colonization resistance against CDI. Despite the extensive knowledge of *C. difficile* virulence, little is known regarding how this bacterium survives at the host-pathogen interface during infection.

C. difficile induced gastrointestinal inflammation

TcdA and TcdB cause severe damage to intestinal epithelial cells [7]. The resulting loss of epithelial integrity results in disrupted tight junctions, inflammation, and increased oxygenation ensuing a hostile environment at the host-pathogen interface [7-9]. Additionally, the toxin-mediated perforations in the intestinal epithelial layer lead to bleeding in the gut and subsequent translocation of erythrocytes into the gastrointestinal lumen [10]. Hemolysis due to pathophysiological stress occurs, resulting in the release of hemoglobin-bound heme and free heme at the site of infection (Figure 1) [11]. Furthermore, the toxins stimulate the production of proinflammatory cytokines and chemokines by resident immune cells and intoxicated epithelial cells initiating the recruitment of circulating innate and adaptive immune cells. Recruited neutrophils, a key characteristic of the clinical pathophysiology of CDI, as well as other immune cells, produce antimicrobial peptides, reactive oxygen species (ROS), and reactive nitrogen species (RNS) at the site of infection in an attempt to restrict the proliferation of *C. difficile* [4, 12]. Despite these stressors, *C. difficile* is able to survive and thrive in the inflamed colon and disease manifestations can progress leading ultimately to death [3].

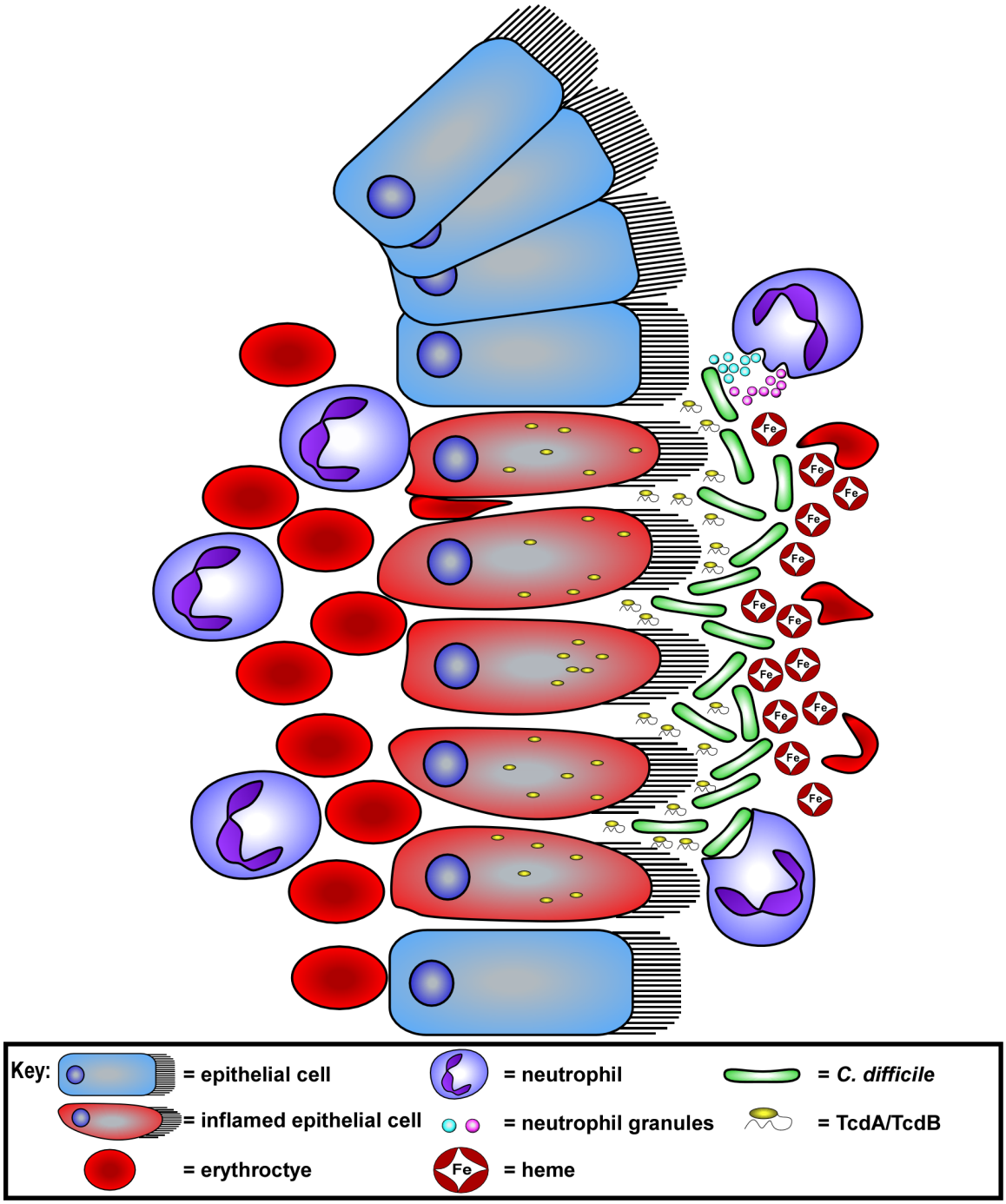

Figure 1 *C. difficile* induced inflammation of the gastrointestinal epithelium. *C. difficile* infects the colon and produces TcdA and TcdB which damage intestinal epithelium cells leading to inflammation, loosing of tight junctions, and immune cell recruitment. Erythrocytes translocate into the gastrointestinal lumen and subsequently lyse releasing a high concentration of hemoglobin bound and free heme at the site of infection.

The heme paradox: utilization and toxicity

Heme, an iron-containing porphyrin, is the most abundant source of iron in the human body and functions as a redox active cofactor of a multitude of enzymes. The functions of these enzymes range from shuttling electrons in the electron transport chain and those involved in respiration to oxidative stress reduction [13, 14]. Many bacteria satisfy the need of heme through regulated biosynthesis, however some bacteria such as *Enterococcus faecalis* and *Lactococcus lactis* strictly utilize exogenous heme as these species do not contain complete heme biosynthetic pathways [14, 15]. Many invading pathogens have evolved heme acquisition mechanisms ranging from surface receptors to secreted proteins to utilize this rich metabolic resource [16-20]. Thus, heme is a sought-after host factor that can be used as a nutritional source and metabolic cofactor in the host environment.

Owing to its reactive nature, heme is toxic to bacteria at high concentrations through a variety of mechanisms, including generation of ROS through Fenton chemistry and damaging membrane proteins and lipids due to the hydrophobic structure of heme. [21-24]. However, a complete understanding of heme-mediated toxicity has not been defined, particularly when extended to toxicity in an anaerobic environment [23]. To defend against the stresses of heme-mediated damage, bacteria encode systems for heme sensing and detoxification [25-31]. In the Gram-positive pathogens *Staphylococcus aureus* and *Bacillus anthracis,* the heme stress response is controlled by the heme sensing two-component system, HssRS, which regulates transcription of the ABC transporter HrtAB to reduce heme toxicity through efflux [25, 26, 32]. Reducing intracellular heme levels through export is a conserved microbial strategy as heme efflux systems have also been identified in *Lactococcus lactis, Streptococcus agalactiae,* and *Neisseria gonorrhoeae* [27, 28, 33]. In each example, inactivation of heme detoxification machinery

increases heme sensitivity and modulates virulence [25-28, 33]. Additionally, the reduction of free heme through protein utilization or sequestration is a strategy employed primarily by Gram-negative pathogens [30, 31, 34]. While the proteins in the conserved HemS family identified in *Yersinia Pestis, Yersinia enterocolitica, Pseudomonas aeruginosa, Escherichia coli,* and *Shigella dysenteriae* contain a variety of heme storage, transfer, or utilization properties, the reduction of free heme protects against toxicity. A final mechanism of heme detoxification consists of cleaving the porphyrin ring to release the reactive iron center by catabolic heme oxygenases, such as IsdI and IsdG in *S. aureus* [35, 36]. The diverse mechanisms of heme detoxification underscore the evolutionary importance of maintaining heme homeostasis in order for human pathogens to cause disease (Figure 2).

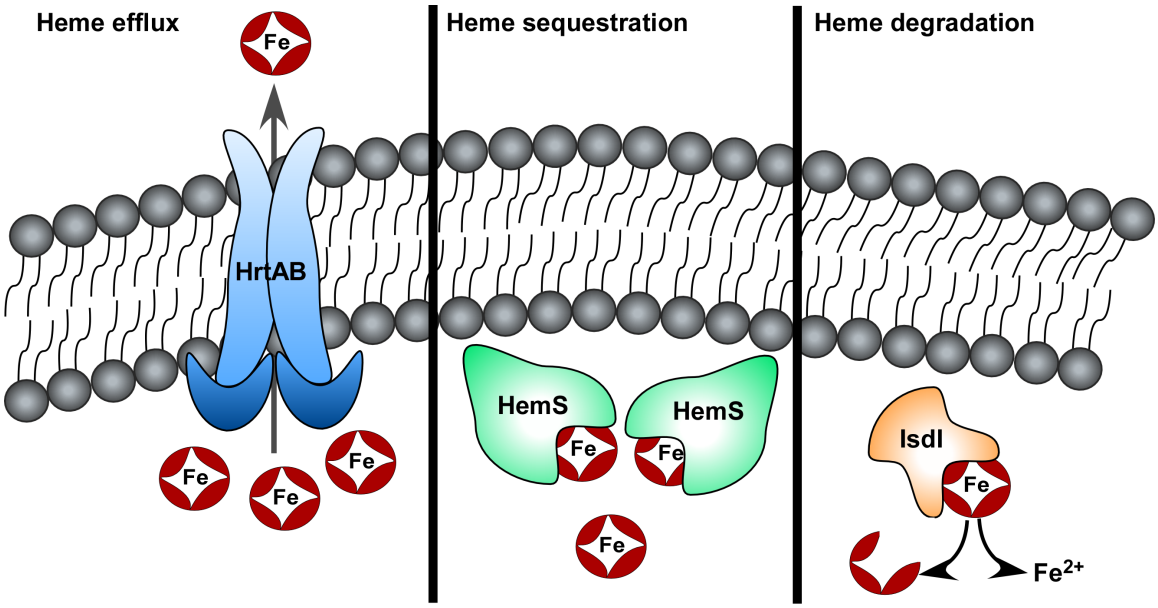

Figure 2 Bacterial mechanisms to reduce heme toxicity. Heme toxicity is alleviated through removal of heme through efflux, sequestration or utilization by heme binding proteins, or degrading heme to the cleaved porphyrin and free iron.

C. difficile and heme

The ability of *C. difficile* to utilize heme has not been defined. *C. difficile* lacks a complete heme biosynthesis pathway, specifically the enzymes required to produce the porphyrin precursor δ-aminolevulinic acid and the iron-inserting enzyme ferrochelatase [20]. However, *C. difficile* contains the biosynthetic enzymes for siroheme and a cobalt containing tetrapyrrole [37, 38]. While a recent study identified an oxygen independent heme degrading enzyme in *E. coli, C. difficile* does not appear to be able to degrade heme as it cannot use heme as a sole iron source [39]. Additionally, *C. difficile* does not respire and no enzymes that require heme as a cofactor such as catalase or other proteins involved in protection against oxidative stress have been identified. These data suggest *C. difficile* either contains evolutionary distinct heme utilization proteins or has evolved without the requirement for heme to survive.

Toxin-mediated release of heme from the gastrointestinal lumen may reach lethal concentrations at the site of CDI. Clostridial species are sensitive to heme toxicity in anaerobic environments despite the mechanism of toxicity remaining undefined [23]. It is also unknown if heme is toxic to *C. difficile* or if this bacterium encounters heme toxicity during infection. Notably, *C. difficile* does not contain orthologs of known heme detoxification systems. The inflamed gastrointestinal tract is a complex environment and *C. difficile* occupies different metabolic niches during infection [40]. This implies that this pathogen may have not evolved molecular mechanisms to reduce heme toxicity, rather it modifies its pathogenesis or physical location during infection to avoid areas of concentrated heme. Ultimately, an understanding of the interplay between *C. difficile* and heme at the host-pathogen interface is severely lacking.

Conclusions

C. difficile toxin-mediated inflammation of the gastrointestinal tract leads to a plethora of stressors that this bacterium must mitigate to maintain full pathogenicity [9]. My dissertation aims to reveal the threat of heme toxicity to *C. difficile* during infection and the mechanisms this pathogen utilizes to survive the harsh environment of the host. The overarching hypothesis of my work states that *C. difficile* encodes mechanisms to sense and persist in high heme concentrations at the site of infection to achieve full virulence. In the subsequent chapters, I visualize a high concentration of hemoglobin in the infected cecum of a mouse and therefore heme at the host-pathogen interface. I discover that heme is toxic to *C. difficile* and this bacterium can adapt to heme toxicity if pretreated with a low concentration of heme. I identify and name the heme activated transporter (*hatRT*) operon that encodes a molecular mechanism to sense and detoxify intracellular concentrations of heme through efflux. Strains that lack the HatT transporter have a reduction in pathogenicity in a toxin independent manner. I further identify and name the heme sensing membrane protein (*hsmRA*) operon that encodes proteins that detoxify heme through sequestration and utilize the bound heme to function as a shield against redox active molecules produced by the host. I reveal the complete transcriptional response in *C. difficile* to a brief exposure of heme to be limited to the *hatRT* and *hsmRA* operons. Strains lacking HsmR or HsmA displayed increased sensitivity to vancomycin and a delay in recovery in a relapse mouse model of infection. Collectively, the results herein describe two mechanisms of heme detoxification and utilization employed by *C. difficile* to survive and cause full disease in the host. These findings set the stage for the development of therapeutic interventions to target these bacterial-specific systems.

A version of the following section (*Chapter II,* HEME SENSING AND DETOXIFICATION BY HATRT CONTRIBUTES TO PATHOGENESIS DURING *CLOSTRIDIOIDES DIFFICLE* INFECTION) was originally published in *PLoS Pathogens* (December 2018).

Knippel, R. J.: Zackular, Z. P.: Moore, J. L.: Celis, A. I.: Weiss, A.: Washington, M. K.: DuBois, J. L.: Skaar, E. P. Heme Sensing and Detoxification by HatRT Contributes to Pathogenesis during *Clostridium difficile* Infection. *PLoS Pathogens* 2018, *14, 12,* e1007486.

doi: 10.1371/journal.ppat.1007486

CHAPTER II

HEME SENSING AND DETOXIFICATION BY HATRT CONTRIBUTES TO PATHOGENESIS DURING *CLOSTRIDIOIDES DIFFICLE* INFECTION

Introduction

The overall goal of this chapter was to investigate the occurrence of heme exposure to *C. difficile* within the gastrointestinal lumen during infection. Here, I visualize increased abundance of hemoglobin in the gastrointestinal lumen as a result of CDI using imaging mass spectrometry. A heme-inducible operon was identified that contains a TetR family transcriptional regulator and major-facilitator superfamily transporter. I have named these gene products HatRT for heme activated transporter (R=regulator, T=transporter). The transcriptional regulator HatR responds to intracellular heme concentrations through binding of heme leading to the de-repression and increased transcription of *hatRT*. Lack of the HatT transporter results in increased intracellular heme concentrations and a decrease in pathogenicity in a murine model of infection. Taken together, these results describe a mechanism by which *C. difficile* detoxifies heme and establishes a requirement for heme sensing and detoxification for full virulence during *C. difficile* infection.

Materials and methods

Ethics statement

All animal experiments under protocol M1700053 were reviewed and approved by the Institutional Animal Care and Use Committee of Vanderbilt University. Procedures were performed according

to the institutional policies, Animal Welfare Act, NIH guidelines, and American Veterinary Medical Association guidelines on euthanasia.

Bacterial strains, growth conditions, and plasmids

Strains used in this study are listed in Table 1. *C. difficile* strains were grown at 37 °C in an anaerobic chamber (85% nitrogen, 10% hydrogen, 5% carbon dioxide, Coy Lab Products) in brain-heart-infusion broth (BD Life Sciences) supplemented with 0.5% yeast extract (BD Life Sciences) and 0.1% cysteine (Sigma-Aldrich) (BHIS) or in *C. difficile* minimal media (CDMM) described previously [41]. *Escherichia coli* strains were grown in lysogeny broth (LB) or agar (LBA), supplemented with 50 μg/mL kanamycin or 50 μg/mL carbenicillin when necessary [41]. *Bacillus subtilis* strains were grown on LBA or in BHI broth supplemented with 5 μg/mL tetracycline or 2.5 μg/mL chloramphenicol. All antibiotics were purchased from Sigma-Aldrich.

hatR::CT and hatT::CT strain generation. Gene inactivations were achieved using the ClosTron system as described previously [42]. Briefly, gBlocks containing specific modifications for insertion into the genome were generated using the TargeTronics algorithm (http://www.targetrons.com) and synthesized by Integrated DNA Technologies. The gBlocks were cloned into pCR-Blunt vector using the Zero Blunt PCR cloning kit (ThermoFisher Scientific) followed by restriction digest with BsrgI and HindIII (NEB) and ligation (NEB T4 ligase) into pJS107. Plasmids were transformed into the *recA*$^+$ *E. coli* MG1655 through a standard heat shock protocol followed by transformation into *B. subtilis* JH2 using an established method [42]. *B. subtilis* strains containing the pJS107*_hatR* or pJS107*_hatT* plasmids were mated with *C. difficile* R20291 overnight at 37 °C by plating and mixing together 100 μL of each strain onto a BHIS plate in the anaerobic chamber. Plates were scraped and transferred into 2 mL of BHIS prior to plating

200 µL onto BHIS plates containing 20 µg/mL thiamphenicol and 50 µg/mL kanamycin ($BHIS_{thiamp20kan50}$). Colonies from these plates were patched onto new $BHIS_{thiamp20kan50}$ and BHIS plates containing 5 µg/mL tetracycline ($BHIS_{tet5}$). Patched colonies that were tetracycline sensitive were patched again onto new $BHIS_{thiamp20kan50}$ and $BHIS_{tet5}$ plates. Colonies that remained tetracycline sensitive were streaked onto BHIS plates containing 20 µg/mL lincomycin ($BHIS_{linc20}$). Inactivation of the *hatR* or *hatT* gene was confirmed by performing PCR to identify a 1.5 kbp shift in gene size using gDNA extracted as previously described on colonies that were lincomycin resistant [42].

xylE reporter and complementation plasmids. Reporter and complementation plasmids (Table 1) were created by GenScript using the pJS116 plasmid as a backbone for the synthesized intergenic region (236 bp) of *hatR* fused to the *xylE* reporter gene, the intergenic and full coding region of *hatR,* and intergenic region of *hatR* fused to the full coding region of *hatT. C. difficile* strains were transformed as described above with the removal of the lincomycin selection and were maintained on $BHIS_{thiamp20}$ to ensure plasmid retention.

Protein expression plasmids. Protein expression plasmids for HatR were generated by amplifying *hatR* flanked by BamHI and XhoI and cloning into the multiple cloning site of pLM302 after restriction digest. Point mutant generation in pLM302_*hatR* was performed with NEB Q5 Site Directed Mutagenesis kit according to the manufacturer's instructions, using the primers listed in Table 2. Mutations to Ala or Leu were governed by the surrounding protein motifs and retention of spatial arrangement.

Heme toxicity growth assays. Freshly streaked bacterial colonies were used to inoculate 5 mL of BHIS or $BHIS_{thiamp20}$ and grown for 16 h at 37 °C. Cultures were subcultured 1:50 into fresh BHIS or $BHIS_{thiamp20}$ and grown for 6 h at 37 °C prior to 1:50 inoculation into CDMM or $CDMM_{thiamp20}$

containing heme at the indicated concentrations. All growth assays were performed in a 96-well plate in 200 μL of media. Optical density at 600 nm (OD_{600}) served as measurement of growth and was measured every 30 min for the indicated total time in an EpochII microplate reader (BioTek).

Table 1 Bacterial strains and plasmids used in chapter II.

Bacterial Strain or plasmid	Relevant Feature or Genotype	Reference
Clostridioides difficile R20291		[43]
Clostridioides difficile hatR::CT	Intron inserted into *hatR*	This chapter
Clostridioides difficile hatT::CT	Intron inserted into *hatT*	This chapter
Bacillus subtilis JH BS2	Carries Tn196	[42]
Escherichia coli DH5α		[44]
Escherichia coli MG1655	RecA+	[45]
Escherichia coli BL21(DE3)		[46]
pJS107	ClosTron plasmid	[42]
pJS107*_hatR*	ClosTron plasmid with intron targeted to *hatR*	This chapter
pJS107*_hatT*	ClosTron plasmid with intron targeted to *hatT*	This chapter
pJS116	Stable *C. difficile* plasmid	[42]
pJS116*_phatR-hatR*	*hatR::CT* complementation plasmid	This chapter
pJS116*_phatR-hatT*	*hatT::CT* complementation plasmid	This chapter
pJS116*_phatR-xylE*	XylE reporter gene driven by the promoter of *hatR*	This chapter
pLM302	Protein expression plasmid	Center for Structural Biology, Vanderbilt University
pLM302*_hatR*	HatR expression plasmid	This study
pLM302*_hatR*-H99L	HatR-H99L expression plasmid	This study

Table 2 Oligonucleotides used in chapter II.

Name	Sequence (5' – 3')	Description
qRT_*hatR*_F	ggaaaatagctcaaaggtgtgg	qRT-PCR forward primer for *hatR*
qRT_*hatR*_R	cgccttatcaataaatcccattc	qRT-PCR reverse primer for *hatR*
qRT_*hatT*_F	ttggaaagccctcaactcc	qRT-PCR forward primer for *hatT*
qRT_*hatT*_R	ttttgccatttcagggtctg	qRT-PCR reverse primer for *hatT*
qRT_*rpoB*_F	tgctgttgaaatggttcctg	qRT-PCR housekeeping gene forward primer
qRT_*rpoB*_R	cggttggcatcatcattttc	qRT-PCR housekeeping gene reverse primer
R20291_*hatR*_F	aggtgtggtataagtgcaggt	Forward primer to check for intron insertion into *hatR*
R20291_*hatR*_R	agctgttcatgaaagtcgtc	Reverse primer to check for intron insertion into *hatR*
R20291_*hatT*_F	gtggtgtttaccttgaatcataat	Forward primer to check for intron insertion into *hatT*
R20291_*hatT*_R	cttgaacctaaaatattggcaatacc	Reverse primer to check for intron insertion into *hatT*
*hatR*_BamHI_F	ggggatccgggatgccaaagattttagaaaatg	Forward primer for cloning *hatR* into pLM302
*hatR*_XhoI_R	gggctcgaggggttaatgtattaatttttcaata	Reverse primer for cloning *hatR* into pLM302
*hatR*_H99L_F	gaaattttataaacaaagttcttaatatccaatacaatg	Forward primer for H99L point mutation in *hatR*
*hatR*_H99L_R	cattgtattggatattaagaactttgtttataaaatttc	Reverse primer for H99L point mutation in *hatR*

RNA extraction and sequencing

C. difficile were grown anaerobically in triplicate in CDMM in 0 or 50 μM heme. Hemin (Sigma) was solubilized in 0.1 M NaOH. The cultures were grown at 37 °C to an OD_{600} of 0.3 abs. Upon reaching this density, a 1:1 solution of acetone:ethanol was added to an equal volume of the culture. Samples were stored at -80 °C until used for RNA extraction. Samples were thawed on ice, pelleted, and resuspended in 750 μL of LETS buffer (1 M LiCl, 0.5 M EDTA, 1 M Tris pH7.4). Cells were transferred to tubes containing lysing matrix B beads (MP Biomedicals) and lysed by a FastPrep-24 (MP Biomedicals) bead beater for 45 s at 6 m/s. Lysed samples were heated for 5 min at 55 °C and pelleted by centrifugation for 10 min. The supernatant was transferred to a fresh tube and 1 mL TRIzol (Thermo Scientific) was added. Chloroform (200 μL) was added to each sample and vortexed prior to separation of the aqueous and organic layers by centrifugation for 15 min. The aqueous (upper) layer was transferred to a fresh tube and the RNA was precipitated through the addition of 1 mL isopropyl alcohol. Samples were incubated for 10 min and RNA was pelleted by centrifugation for 10 min. Supernatant was removed and the RNA pellet was washed with 200 μL of 70% ethanol. Samples were air dried for 1 min, then resuspended in 100 μL RNase free water. DNA contamination was removed through the addition of 8 μL RQ1 DNase, 12 μL 10x RQ1 buffer, and 2 μL RNase inhibitor (Promega) to the purified RNA. Samples were DNase treated for 2 h and purified using the RNeasy miniprep RNA cleanup kit (Qiagen). RNA concentration was determined using the Synergy 2 with Gen 5 software (BioTek).

RNA-seq library preparation and sequencing. RNA-seq library construction and sequencing was performed by HudsonAlpha. Concentration was determined using the Quant-iT RiboGreen RNA assay (Thermo Scientific) and integrity was visualized using an RNA 6000 nano chip (Agilent) on an Agilent 2100 Bioanalyzer (Applied Biosystems). RNA was normalized to 500 ng of total RNA

for each sample and the ribosomal RNA (rRNA) was removed using Ribo-Zero rRNA Removal Kit (Illumina). Directly after rRNA removal, the RNA was fragmented and primed for first strand synthesis using the NEBNext First Strand synthesis module (New England BioLabs Inc.) followed by second strand synthesis using NEBNext Ultra Directional Second Strand synthesis kit. Library preparation was achieved using NEBNext DNA Library Prep Master Mix set for Illumina with minor modifications. PolyA addition and custom adapter ligation was performed following end – repair. Post-ligated samples were individually barcoded with unique in-house Genomic Services Lab primers and amplified through 12 cycles of PCR. Library quantity was assessed by Qubit 2.0 Fluorometer (Invitrogen), and quality was determined using a DNA High Sense chip on a Caliper Gx (Perkin Elmer). Final quantification of the complete libraries for sequencing applications was measured using the qPCR-based KAPA Biosystems Library Quantification kit (Kapa Biosystems, Inc.). Libraries were diluted to 12.5 nM and pooled equimolar prior to clustering. Paired-End sequencing was performed on an Illumina HiSeq2500 sequencer (Illumina, Inc.). Raw sequence data are deposited on the NCBI Sequence Read Archive.

Processing of RNA-seq reads. RNA-seq analysis was performed by HudsonAlpha utilizing their unique in-house pipeline. Briefly, quality control was performed on raw sequence data from each sample using FastQC (Babraham Bioinformatics). Curated raw reads were imported into the data analysis platform, Avadis NGS (Strand Scientifics) and mapped to the reference *C. difficile* R20291 genome. Aligned reads were filtered on various criteria to ensure the highest read quality. Replicate samples were grouped and quantification of transcript performed using Trimmed Means of M-values as the normalization method. Differential expression of genes was calculated using fold change (using default cut-off $\geq \pm 2.0$) observed between conditions, and the p-value of the differentially expressed gene list was estimated by Z-score calculations using determined by

Benjamini Hochberg FDR correction of 0.05 [47]. The genome alignment figure (Fig 2A) was created using Circos with a max of 30,000 RPKM displayed.

Quantitative RT-PCR

RNA was extracted as described above and 2 μg was reverse transcribed by M-MLV reverse transcriptase (Fisher Scientific) in the presence of RNase inhibitor (Promega) and random hexamers (Promega). Reactions lacking the reverse transcriptase were used to control for DNA contamination. Newly created cDNA was diluted 1:100 and was used in qRT-PCR using iQ SYBR green supermix (BIO-RAD) utilizing the primer pairs in Table 2. Amplification was achieved using a 3-step melt cure program on a CFX96 qPCR cycler (BIO-RAD). Transcript abundance was calculated using the ΔΔCT method normalized by the *rpoB* gene.

Polyclonal antibody generation

HatR was purified as described below and fresh protein was submitted to the Vanderbilt Antibody and Protein Resource core for generation of a rabbit polyclonal antibody against HatR. This antibody was affinity purified for increased HatR specificity. The α-HatR antibody was tested for specificity and reactivity in immunoblot analysis of purified HatR protein in addition to whole cell lysates from heme treated WT and *hatR::CT* strains.

Immunoblotting analysis

WT or *hatR::CT* strains were grown in 5 mL of BHIS overnight at 37 °C. Cultures were subcultured into fresh BHIS containing 0, 1, 5, 10 or 25 μM heme and grown for 6 h. Cells were pelleted by centrifugation (4000 x *g* for 10 min), supernatant was removed and were resuspended

in 1 mL of 1 X PBS containing 2.5 mg/mL lysozyme (ThermoFisher Scientific). Samples were incubated for 1 h at 37 °C, pelleted by centrifugation (20,000 x *g* for 5 min), then resuspended in 1 X PBS followed by sonication using Ultrasonic dismembrator (ThermoFisher Scientific) to lyse the cells. Debris from the lysed cells was pelleted by centrifugation (20,000 x *g* for 5 min). Supernatant was used in immunoblotting analysis using rabbit polyclonal anti-HatR antibodies as previously described [48]. Detection was performed using a goat anti-rabbit IgG (H+L) cross-adsorbed secondary antibody with an Alexa Fluor 680 and imaged using a ChemiDoc MP imaging system (Bio-Rad).

Protein expression and purification

E. coli BL21 (DE3) pREL containing the pML302_*hatR* plasmids were grown overnight in 5 mL of LB_{kan50} at 37 °C. Cells were subcultured into Terrific broth (ThermoFisher Scientific) containing 50 µg/mL kanamycin and grown to the mid-logarithmic phase of growth (0.5 abs measured at 600 nm) at 37 °C prior the addition of 1 mM isopropyl-1-thiol-D-galactopyranoside (IPTG). Growth was continued at 16 °C for 16 h. Cells were harvested by centrifugation (6000 x *g* for 10 min) and resuspended in 1 X PBS. Cells were lysed by passage through an EmulsiFlex homogenizer (Avestin) three times at 20,000 lb/in^2. The insoluble debris was removed by centrifugation at 40,000 x *g* for 1 h and the supernatant was filtered using a 0.22-µM-pore sizer filter. Filtered lysate was added to amylose resin (New England Biolabs Inc.) and allowed to bind at 4 °C for 30 min prior to transfer to a gravity column. The column was washed with four column volumes of wash buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, pH 7.5) three times followed by 2 column volumes of elution buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 10 mM maltose, pH 7.5) twice. The maltose-binding protein tag was cleaved using the Pierce HRV 3C Protease

Solution kit (ThermoFisher) by following the manufacturer's instructions. Cleaved tag and protease was removed by the addition of HisPur Cobalt Resin (ThermoFisher) and allowed to bind at 4 °C for 1 h with rotation. Beads were pelleted by centrifugation (2000 x *g* for 2 min) and the supernatant containing tagless protein was removed.

Absorption spectroscopy

Heme binding by HatR were determined by measuring the absorption spectrum of increasing amounts of hemin (0 – 25 μM) after addition to a cuvette containing 10 μM recombinant HatR in 1 mL of Tris-buffered saline (TBS) and a reference standard containing 1 mL TBS on a Varian Cary 50BIO. Samples were mixed and allowed to incubate at room temperature in the dark for 5 min prior to collecting the spectrum between 300 – 800 nm with 10 nm increments. Binding ratio of heme to HatR was determined by plotting the change in absorbance at 413 nm between the reference standard and the HatR sample. A curve fit and ratio was obtained by performing the one-site binding model non-linear regression function on Graph Pad Prism 6.

XylE reporter assays

Bacteria harboring the reporter plasmid pJS116_*phatR-xylE* were grown overnight in $BHIS_{thiamp20}$ and subcultured 1:50 into 10 mLs of fresh $BHIS_{thiamp20}$ containing 0 or 10 μM heme. Cultures were grown for 6 h at 37 °C prior to cytoplasmic fraction preparation and analysis of XylE activity as described previously [26]. Absolute XylE activities were determined spectrophotometrically by measuring the formation of 2-hydroxymoconic acid from catechol for *C. difficile* reporters due to lysozyme interference during protein quantification.

LC-MS heme quantification

Bacterial growth. C. difficile WT and *hatT::CT* strains were streaked onto BHIS and grown for 16 h at 37° C. Single colonies were used to inoculate 5 mL cultures in BHIS and grown for 16 h at 37° C. Seven hundred fifty µL of these cultures were subcultured into 75 mL of BHIS and BHIS + 25 µM heme in an Erlenmeyer flask and grown for 16 h at 37° C. Total CFU were determined by serial dilution and plating onto BHIS for enumeration, and cells were collected by centrifugation and flash frozen in liquid nitrogen before storage at -80 °C.

Preparation of standard curves. A 2 mM stock solution of a heme standard was prepared in DMSO. This stock was then diluted to make standards from 0.25-6 µM in acetonitrile + 0.1% trifluoracetic acid (TFA).

Extraction of heme. C. difficile cell pellets were thawed on ice. 1 mL of 1M HCl:DMSO (1:1, v/v) was added and samples were vortexed. Samples were transferred to 2 mL FastPrep lysis B matrix tubes and the cells lysed by bead beating in a FastPep-24 5G instrument (6.0 m/sec, 40 s total, 2X). The cell lysate was centrifuged to pellet debris (10,000 rpm, 5 min, 4°C). Supernatants were collected and kept in the dark. The pellets were resuspended in 1 mL of 1M HCl:DMSO (1:1, v/v), vortexed vigorously, and centrifuged again as above. The supernatants were pooled and the resuspension/centrifugation cycle repeated one more time. The pooled supernatants were filtered using a 0.22 µm Millex-GS syringe filter (MF-Millipore) and subsequently diluted to 25 mL with ddH2O. Using a Sep-Pak Vac 3cc tC18 cartridge (Waters 036815), the extracts were concentrated and subsequently eluted with 2 mL of acetonitrile + 0.1% TFA then 2 mL of methanol. Extracted porphyrins were dried under N2 (g) and resuspended in 100 µL of acetonitrile + 0.1% TFA. All samples were immediately dispensed into vials for mass spec analysis.

LC-MS analysis. Samples were prepared by adding 25 µL of ultrapure water to 75 µL solutions of analytes in acetonitrile + 0.1% TFA. A PLRP-S column (Agilent) was equilibrated to an 85:15 ratio of solvent A (ultrapure water + 0.1% formic acid) to solvent B (acetonitrile + 0.1% formic acid). Liquid chromatography separations were achieved by linear gradient elution, transitioning from 15% to 95% solvent B over 6 min followed by a 2 min hold at 95% B. The column was re-equilibrated to 15% solvent B for 2 min in between injections of the same sample (two technical replicates run per sample, 2 µL injection volume, 600 µL/min flow rate, 50°C). Two blank runs were implemented between samples to ensure against column holdover of analytes. Electrospray ionization mass spectrometry analysis was carried out in positive mode with a capillary voltage of 2Hz (Agilent 6538 UHD Q-TOF).

Quantification of standards and analytes from LC-MS data. Data were analyzed using MassHunter Qualitative Analysis Software and MZmine 2 [49]. Extracted ion chromatograms (EICs) were derived for each individual standard on the basis of its mass per charge (*m/z*) in positive ion mode, which is equivalent to the exact mass of its positive ion $(M+H)^+$. Values for *m/z* were determined empirically for all standards. Peaks associated with each analyte were integrated. For the generation of standard curves, integrated peak areas were plotted versus concentration. Linear regression analysis (Kaleidagraph) was used to determine the correlation coefficient between integrated peak area and heme concentration (slope of standard curve, mporph). For the quantification of analytes from cells, values for the integrated peak intensities (measured in units of ion counts) were converted to units of concentration (µmol/L injected) via: counts x $(mporph)^{-1}$. The concentration of each analyte in the injected volumes [A] was subsequently converted to units of nmol analyte per CFU in sample as: [A] x (volume used to resuspend dried sample) x

(CFUs in analyzed cell pellet)$^{-1}$. Reported values are averages of 3 biological samples and two technical replicates.

Mouse model of CDI

Adult (8 – 12 week old) age-matched male C57Bl/6 (Jackson Laboratories) were housed in groups of five and maintained at Vanderbilt University Medical Center Animal Facilities. Mice were subjected to a previously described model of CDI [50, 51]. Briefly, mice were treated with 0.5 mg/mL cefoperazone in their drinking water for 5 days. Mice were given a 2 day recovery period prior to administration of 10^5 spores of WT, *hatR::CT,* or *hatT::CT C. difficile* strains in PBS via oral gavage. Prior to infection, mice were confirmed to be *C. difficile* negative. After infection, mice were monitored for signs of disease, including diarrhea and weight loss. Mice that displayed severe disease or weight loss greater than 20% were humanely euthanized.

Bacterial burden determination. C. difficile CFUs were quantified daily from fecal samples. Samples were diluted and homogenized in PBS and serial plated onto taurocholate cycloserine cefoxitin fructose agar (TCCFA) for enumeration as CFU per gram of feces.

Histological analysis. On the final day of infection and necropsy, ceca were harvested, fixed in a 10% formalin solution and embedded in paraffin. Cut sections were stained with hematoxylin and eosin (H&E). Stained sections were assigned a disease score in a blinded fashion by a pathologist based on previously established criteria [52]. Histological scores are presented as a sum of three independent criteria: epithelial damage, edema, and inflammation.

Imaging mass spectrometry. MALDI IMS was performed as previously described [53, 54]. Briefly, ceca were harvested after necropsy and flash frozen in liquid nitrogen in a 25% Optimal Cutting Temperature compound. Sections were sequentially washed to remove interfering lipids, salts, and

OCT using 70% ethanol for thirty seconds, 100% ethanol for thirty seconds, 6:2:1 ethanol:chloroform:acetic acid for 2 minutes, 70% ethanol for thirty seconds, and 100% ethanol for thirty seconds. Slides were dried in a desiccator before MALDI matrix was applied. Fifteen mg/mL 2,5-dihydroxyacetophenone was prepared in 90% acetonitrile with 0.2% TFA and crystals were dissolved by sonication for ten minutes. Matrix was applied six times using a TM-Sprayer (HTX Imaging) operated at 1100 mm/min and at a flow rate of 0.2 mL/min using 90% acetonitrile as a pushing solvent. The spray nozzle was heated to 80 °C with the track spacing set to 2 mm. Coating was rehydrated using 1 mL of 50 mM acetic acid in a sealed hydration chamber for 3 min at 85 °C. IMS was performed using a rapifleX MALDI Tissuetyper (Bruker Daltonics) operated in linear positive ion mode. The laser was operated at 10,000 hertz in single mode and pixels were set to be 50 by 50 μm. A total of five-hundred laser shots were captured per pixel with fifty laser shots at each position within the pixel. Data were processed using fleXimaging version 4.1. Data were further analyzed using SCiLS Lab 2015b version 3.02.7774 (Bruker Daltonics). Spectra were normalized to total ion count and baseline subtracted using a top hat algorithm. The images display the ion map of the m/z value of 14,995 without denoising but with interpolation turned on.

C. difficile toxin cytotoxicity determination

Green African monkey kidney epithelial (Vero) cell-rounding cytotoxicity assays were performed as previously described [51]. Cells were grown to confluence in Dulbecco modified Eagle medium (Gibco Laboratories) with 1% penicillin-streptomycin (Gibco Laboratories) and 10% fetal bovine serum (Gibco Laboratories) prior to plating at a total cell density of 10^5 cells per well in a 96-well plate. Fresh fecal samples were normalized to weight, diluted and homogenized in sterile PBS. Fecal debris was pelleted by centrifugation (13,000 g) and tenfold serial dilutions of supernatants

were added to the wells of Vero cells. Complete cell-rounding for each dilution was assessed after overnight incubation at 37 °C with 5% CO_2. Confirmation of *C. difficile* toxin A and toxin B were achieved by neutralization of cell rounding with a combined antitoxin antisera (Techlab). Cell rounding cytotoxicity titers are presented as the $\log_{10}$ of the reciprocal value of the highest dilution with complete rounding of cells.

Statistical Analysis

All data analysis and statistical tests were performed in GraphPad Prism 8 software. Specific statistical tests, replicate numbers, calculated errors and other information for each experiment are reported in the figure legends.

Results

***C. difficile* infection increases hemoglobin abundance in the gastrointestinal lumen.**

To identify host proteins that increase in abundance during *C. difficile* infection (CDI), I applied matrix-assisted laser desorption ionization imaging mass spectrometry (MALDI IMS) to a murine model that induces susceptibility to infection through administration of cefoperazone (0.5 mg/mL) [50, 51, 53, 54]. Ceca of mice infected with *C. difficile* R20291 presented with high levels of epithelial damage, edema, and inflammation on day 4 of the infection (Figure 3A). This inflammatory response correlated with a high abundance of the alpha chain of hemoglobin at the sites of pathology and in the luminal space (Figure 3B and 4). In contrast, mouse ceca mock infected with PBS did not exhibit pathology (Figure 3A) and displayed a low abundance of hemoglobin alpha concentrated at the periphery of the intestinal epithelial villi (Figure 3B and 4).

These data demonstrate that CDI leads to high concentrations of hemoglobin at this host-pathogen interface, and considering each hemoglobin protein contains four molecules of heme, support a model whereby *C. difficile* experiences heme stress during infection.

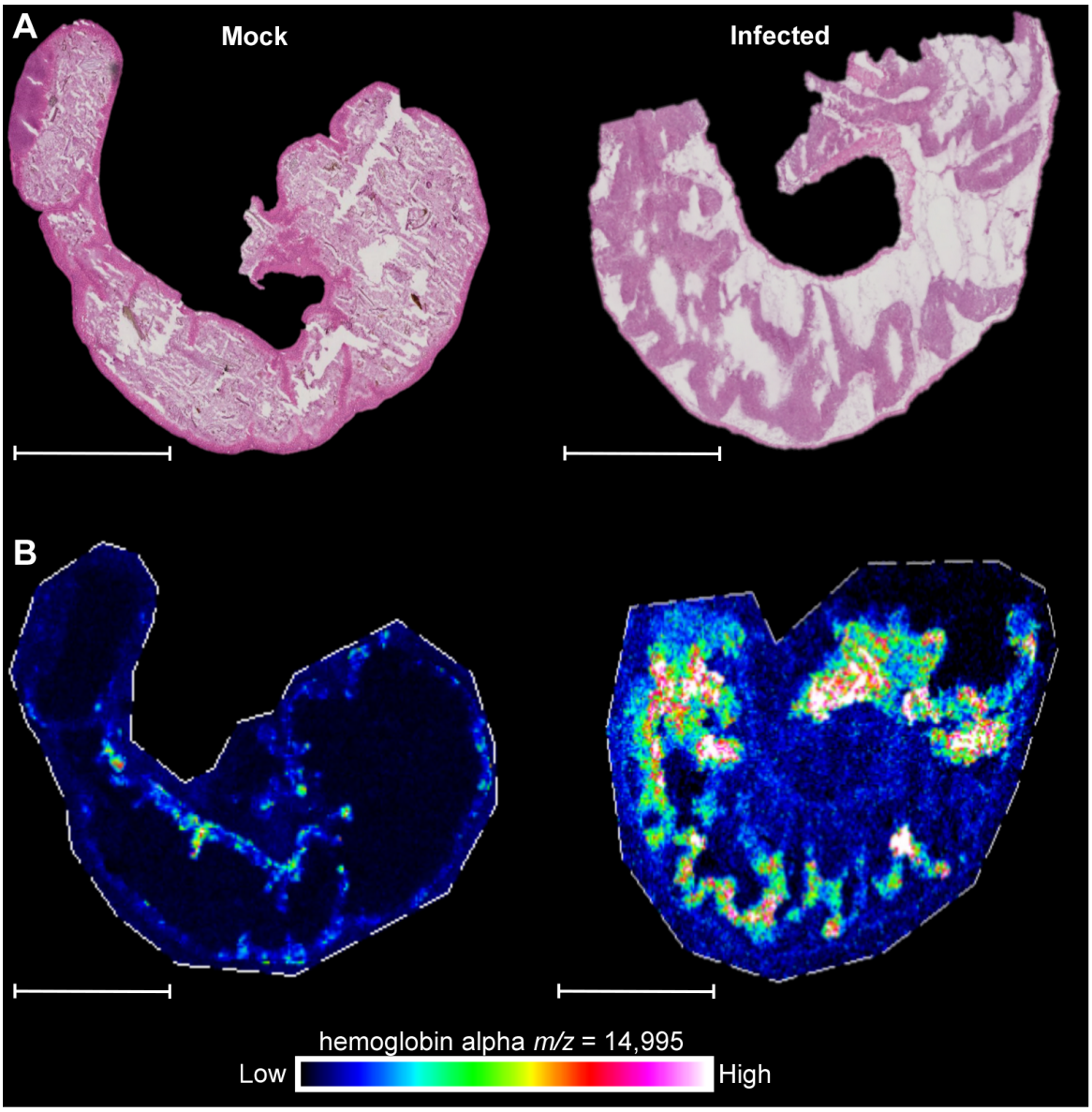

Figure 3 Hemoglobin accumulates in the cecum during *C. difficile* infection. (A) Representative H&E images of mock-infected and *C. difficile* strain R20291 infected C57BL/6 mice. (B) Abundance of hemoglobin subunit alpha in serial sections of the same ceca determine by MALDI IMS. Each image is a representative of 5 independent ceca. Scale bars, 5 mm.

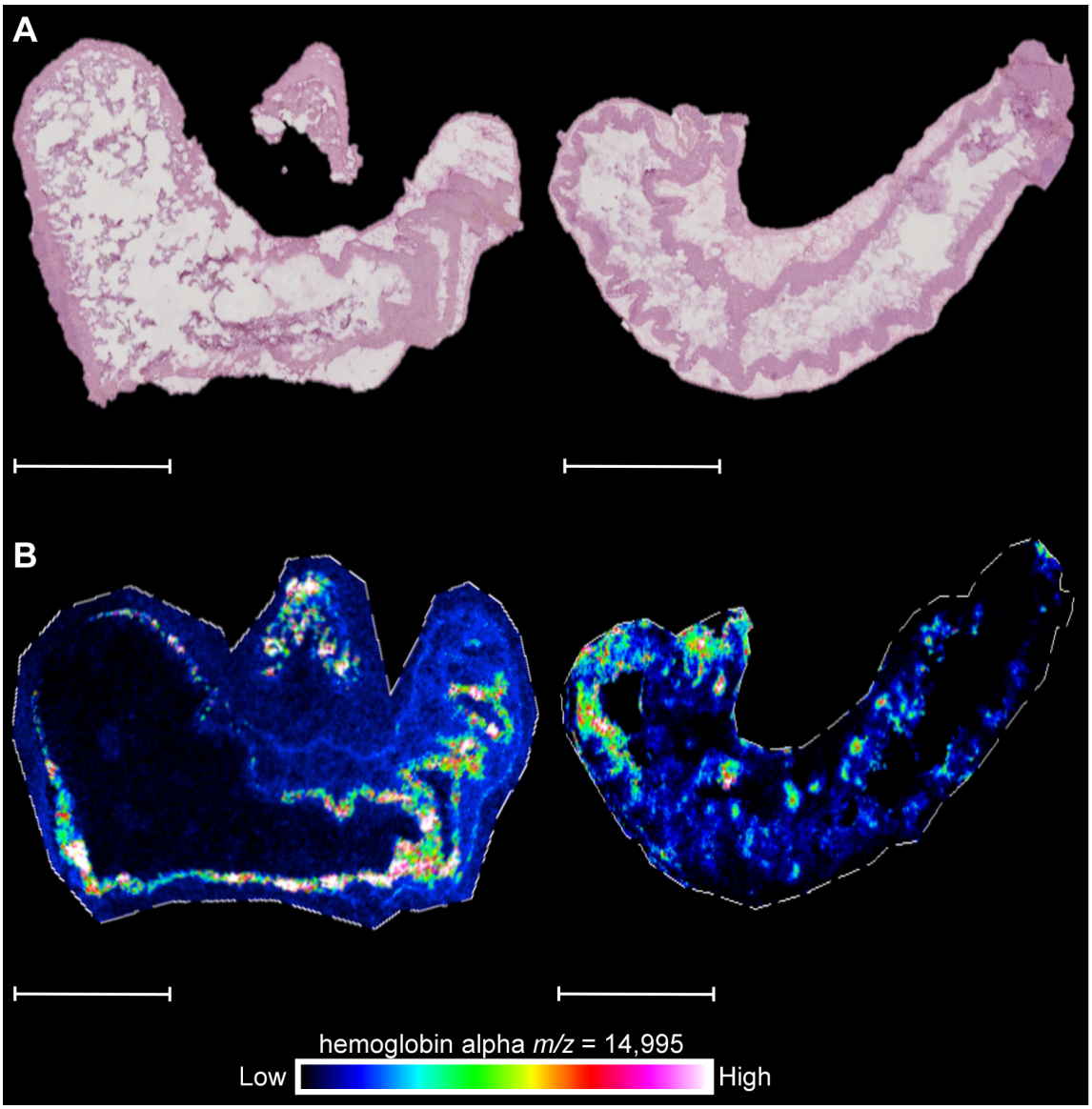

Figure 4 Additional images of hemoglobin accumulation in the cecum during C. difficile infection. (A) H&E images of *C. difficile* strain R20291 infected C57BL/6 mice. (B) Abundance of hemoglobin subunit alpha in serial sections of the same ceca determined by MALDI IMS. Scale bars, 5 mm. Each image pair is an independent cecum from a distinct mouse.

The transcriptional response of *C. difficile* to heme exposure.

Considering the high concentration of hemoglobin in the infected lumen and the reactive nature of heme, we investigated the sensitivity of *C. difficile* to heme toxicity [23]. When *C. difficile* was grown over time in increasing concentrations of heme (0 – 200 μM), a dose-dependent increase in toxicity was observed with a complete inhibition of growth at the highest concentration (Figure 5A). To determine if *C. difficile* can adapt to heme exposure, the growth of *C. difficile* cells pre-exposed to a low concentration of heme (1 μM) was measured following sub-culturing into media containing varying concentrations of heme (0 – 200 μM; Figure 5B). Heme pre-exposure corrected the growth defects of *C. difficile* cultures not pre-exposed to heme (Figure 5A–B), suggesting that *C. difficile* has an inducible mechanism for heme detoxification.

In order to identify the genes that encode proteins responsible for heme adaption, I performed an RNA-sequencing experiment comparing the total relative mRNA transcript abundance of early exponential phase (OD_{600} = 0.3) untreated cultures of *C. difficile* to cultures grown in 50 μM heme. Heme induced the transcription of 245 genes and decreased the transcription of 146 genes (Figure 5C, Appendix A Table 1 and 2). This dataset was curated by grouping significantly upregulated genes that could function as a mechanism of heme sensing and detoxification. Within this group an operon of two genes encoding a TetR family transcriptional regulator (CDR20291_1227) and a major facilitator super family (MFS) transporter (CDR20291_1226) were identified as candidates for further investigation (Figure 6A). These results demonstrate that *C. difficile* has heme responsive genes that may account for its ability to resist and adapt to heme toxicity.

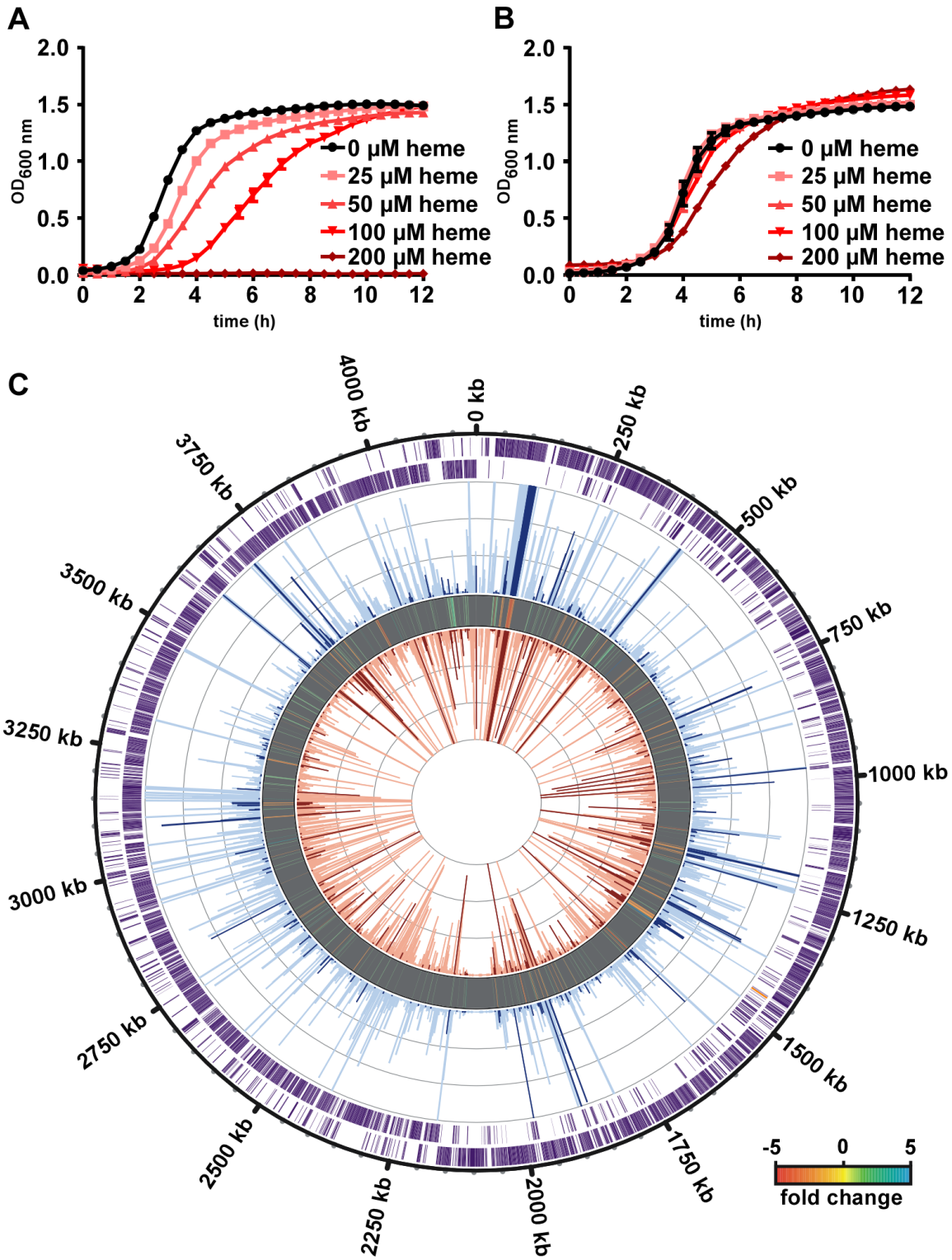

Figure 5 *C. difficile* encodes machinery to detoxify heme. (A) Growth of *C. difficile* R20291 in CDMM containing increasing concentrations of heme. (B). Growth of overnight heme treated (1 μM) *C. difficile* in CDMM containing the exact concentrations of heme as in A. For A and B, the data are the average of means from at least three independent experiments each in biological

triplicate with standard error of the mean shown. (C) RNA-sequencing analysis comparing RNA from heme treated (50 μM, total transcript abundance shown in red peaks inside circle where dark red peaks are significantly changed genes) *C. difficile* to an untreated control (total transcript abundance shown in blue peaks middle circle where dark blue peaks are significantly changed genes). Fold change differences are shown in the circle between the control and heme treated samples with a 5-fold cut off depicted according to the indicated heap map. Outside two rims represent genes in the *C. difficile* R20291 genome (outside is coded in the forward direction, inside is coded in the reverse direction, orange denotes *hatRT* operon).

The *hatRT* operon increases expression in response to heme and confers heme resistance.

To confirm that CDR20291_1227 and CDR20291_1226 are heme responsive, cultures were grown in equimolar concentrations of NaOH (vehicle), protoporphyrin IX (porphyrin ring without iron), iron sulfate, or heme prior to harvesting RNA at the early exponential phase of growth ($OD_{600} = 0.3$). Quantitative reverse transcription PCR (qRT-PCR) was performed on cDNA generated from these samples using primers specific for the genes within this operon. Transcription of both genes were minimally increased in the samples treated with NaOH, protoporphyrin IX, and iron sulfate in contrast to a 2-3 log increase in transcript abundance of the heme treated samples compared to the untreated control (Figure 6B). Due to this considerable transcriptional response to heme, as well as data described below, we named CDR20291_1227 heme activated transporter regulator (*hatR*) and CDR20291_1226 heme activated transporter (*hatT*).

To investigate the heme responsive abundance of HatR, polyclonal antiserum was generated against recombinant HatR and immunoblot analyses were performed on whole cell lysates grown in increasing concentrations of heme. The increase in HatR protein abundance correlated with the increase in concentration of heme, further supporting the observation that *hatRT* is up-regulated upon heme exposure (Figure 6C). To demonstrate the specificity of this antisera, we generated a strain of *C. difficile* inactivated for *hatR* (*hatR::CT*) using the ClosTron system [55]. In this strain, HatR is no longer produced in response to heme (Fig 3C). The lack of *hatR* renders the bacteria more sensitive to heme toxicity, as growth over time in the presence of 50 μM heme is delayed in the mutant compared to wild-type (WT, Figure 6D). A more significant growth delay is observed when *hatT* is inactivated (*hatT::CT*) using the ClosTron system and exposed to the same concentration of heme (Figure 6D). The growth of *hatR::CT* and *hatT::CT* strains are restored to WT levels by expressing *hatR* or *hatT,* respectively, *in trans* under the

control of the intergenic region upstream of *hatR* (Figure 7). Together these data suggest that HatR and HatT coordinate to sense, respond to, and alleviate heme toxicity.

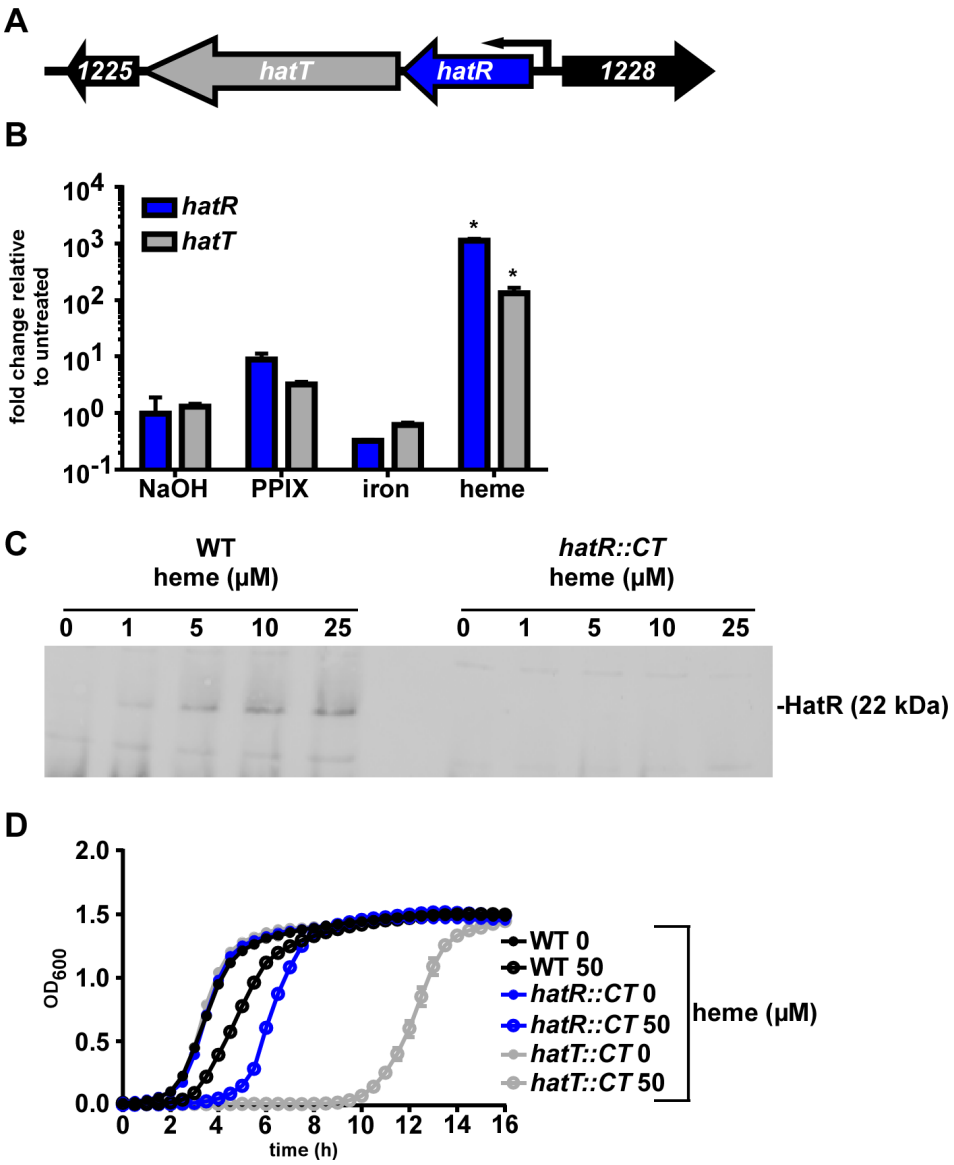

Figure 6 The *hatRT* operon responds to and relieves heme toxicity. (A) Schematic of the *hatRT* operon. (B) *hatR* and *hatT* transcription determined by qRT-PCR. cDNA was reverse transcribed from RNA harvested from *C. difficile* R20291 grown in the presence of sodium hydroxide (NaOH, 500 μM), protoporphyrin IX (PPIX, 50 μM), iron sulfate (50 μM) or heme (50 μM). Transcription is graphed as the fold change relative to an untreated control. The data are a representative of three independent experiments each in biological triplicate with standard deviation shown. Statistical significance was determined using the multiple comparison one-way ANOVA test comparing the means of each group to one another * denotes $p < 0.001$ (C) Immunoblot for HatR from *C. difficile* WT and *hatR::CT* whole cell lysates grown in the presence of increasing concentrations of heme (0 – 25 μM). Blots are representative of three independent experiments. (D) Growth of WT, *hatR::CT,* and *hatT::CT* strains in the presence or absence of heme (50 μM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

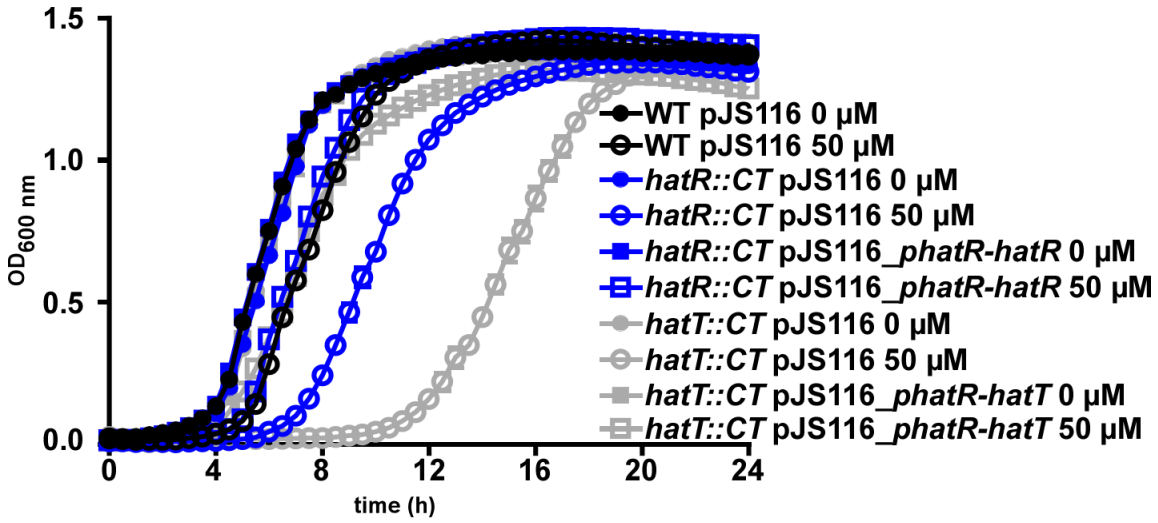

Figure 7 Complementation of *hatR::CT* and *hatT::CT* heme sensitivity. Growth of *C. difficile* WT pJS116 (empty vector), *hatR::CT* pJS116, *hatR::CT* pJS116_*phatR-hatR*, *hatT::CT* pJS116, and *hatT::CT* pJS116_*phatR-hatT* strains in CDMM in the presence or absence of heme (50 µM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean. µM refers to concentration of heme.

HatR functions as a transcriptional repressor of the *hatRT* operon.

As most members of the TetR family of transcriptional regulators directly bind their effector molecules, I examined the ability of HatR to bind heme [56]. Recombinant HatR (10 μM) was incubated with heme (0 – 25 μM), resulting in the appearance of a Soret peak at 413 nm (Figure 8A), indicative of HatR-heme complex formation [57]. Differential absorption spectroscopy at 413 nm over a range of heme concentrations was used to determine that HatR binds heme at a 1:1 ratio using a single site binding model (k_d = 9.2 ± 1.8 μM; Figure 8A insert). To identify the residues responsible for heme binding by HatR, each histidine within HatR was individually mutated. Histidine residues were chosen for substitution due to histidines commonly serving as axial ligands that bind heme [27]. Recombinant proteins containing each individual histidine substitution were purified and heme binding was measured. Substitution of histidine 99 to leucine (H99L) was sufficient to abrogate heme binding (Figure 8B and 9). The substitutions of the remaining four histidines to alanine or leucine (H121A, H126L, H165A, and H180A; Figure 10) did not significantly alter heme binding. These data specify histidine 99 as a critical residue in the formation of the HatR-heme complex.

An examination into the regulation of the *hatRT* operon was performed by creating a plasmid containing a fusion of the intergenic region prior to *hatR* to the reporter gene *xylE,* and transforming this plasmid into WT *C. difficile* [32]. Exposure of this reporter strain to 10 μM heme led to a significant increase in XylE activity as compared to an untreated control (Figure 8C), indicating that heme treatment induces the transcription of the *hatRT* operon in the WT strain. However, upon transformation of the reporter plasmid into the *hatR::CT* strain, there was no significant difference in XylE activity between the untreated or heme exposed samples (Figure 8C). Moreover, the level of XylE activity of the untreated *hatR::CT* strain was significantly higher

than the heme-exposed WT strain, suggesting constitutive expression of *xylE* in the absence of HatR. Taken together, these data suggest that HatR functions as a transcriptional repressor of the *hatRT* operon and that de-repression is achieved through the formation of a HatR-heme complex.

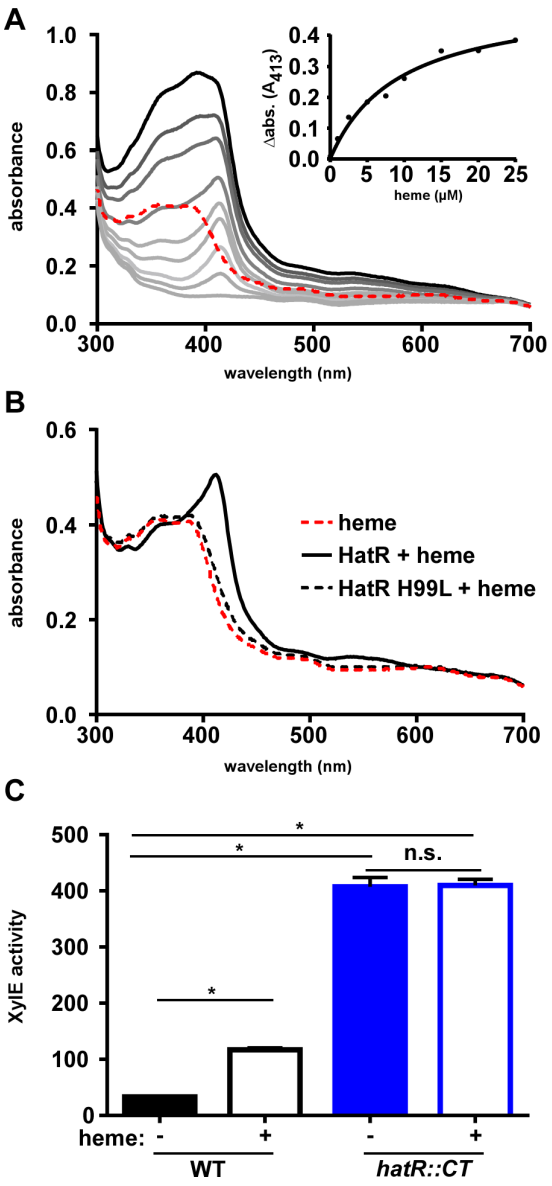

Figure 8 HatR transcriptional repression of the *hatRT* operon is released through direct heme binding. (A) Absorption spectra of heme binding to recombinant HatR. Increasing concentrations of heme (2.5 to 25 μM) were added to 10 μM protein. The spectrum corresponding to 25 μM heme is shown as a dashed red line. HatR with increasing concentrations of heme are shown as gray lines. The inset displays change in absorbance at 413 nm for HatR bound to heme minus the corresponding heme alone peak. (B) Absorption spectra of 10 μM heme binding to HatR and HatR H99L. (C) XylE catechol oxidase activity was measured in *C. difficile* WT and *hatR::CT* strains harboring a *hatR* promoter XylE reporter plasmid after growth in vehicle or 10 μM heme. The data are an average from three independent experiments each in biological triplicate with standard deviation. Statistical significance was determined using the multiple comparison one-way ANOVA test with the Tukey correction for multiple comparisons comparing the means of each group to one another * denotes $p < 0.001$, n.s. denotes not significant.

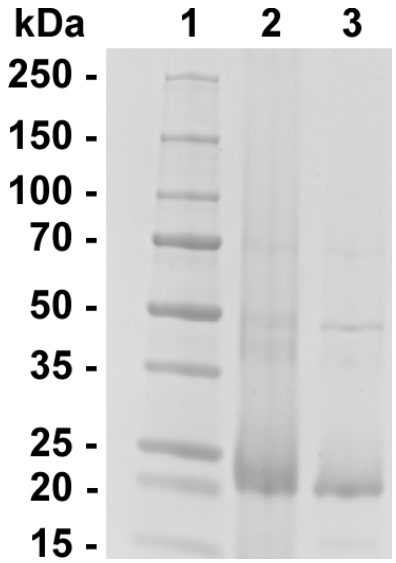

Figure 9 Purified recombinant HatR and HatR H99L. Coomassie stained SDS-PAGE of purified recombinant HatR and HatR H99L. 1 = protein ladder. 2 = HatR (22 kDa). 3 = HatR H99L (22 kDa).

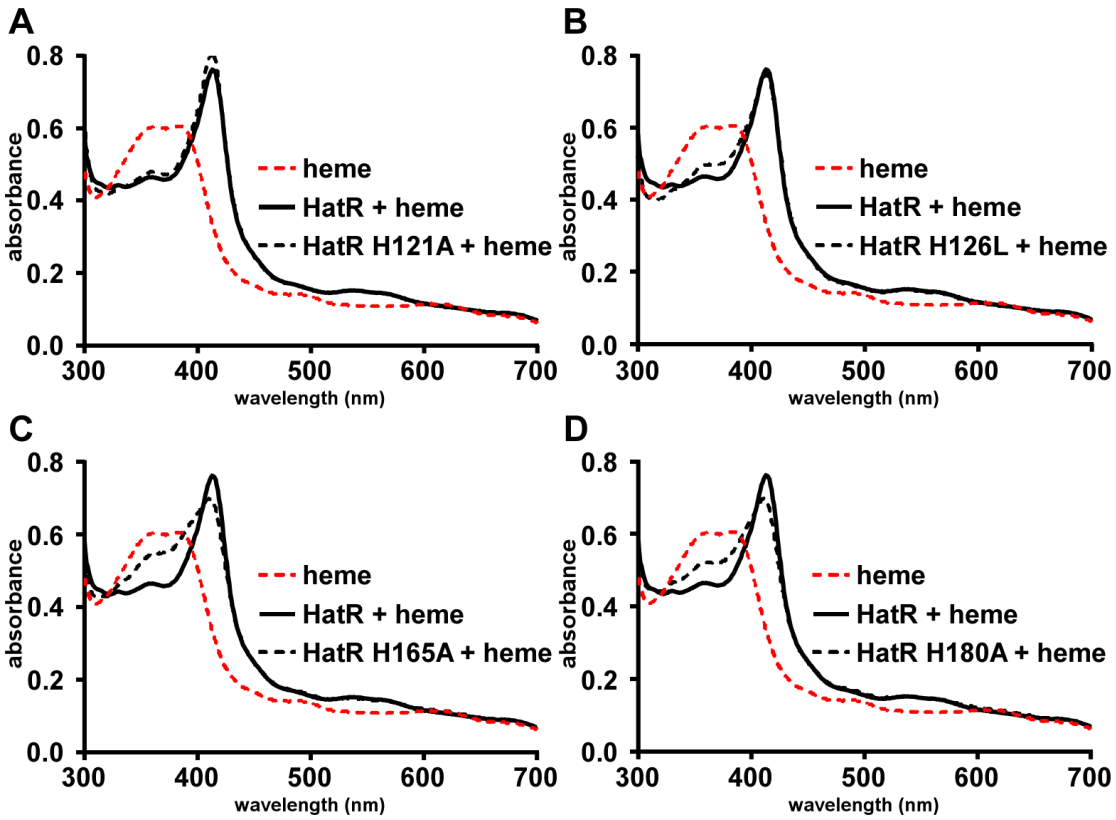

Figure 10 Histidine residues within HatR that are not required for heme binding. Histidine residues within HatR that are not required for heme binding. Absorption spectra of 10 μM heme binding to 10 μM HatR, HatR H121A (A), HatR H126L (B), HatR H165A (C), and H180A (D).

HatT reduces intracellular heme concentrations.

One strategy for microbial heme detoxification involves the reduction of intracellular heme concentrations through efflux [25-28, 33]. To investigate if heme efflux is responsible for HatT-dependent resistance to heme toxicity, I grew the WT and *hatT::CT* strains in the presence or absence of heme (25 μM) for 16 h and measured intracellular heme concentrations utilizing LC-MS analysis. The WT strain treated with heme exhibited a two-log increase in intracellular heme levels when compared to untreated WT cells (Fig 5). In contrast, a more dramatic trend was observed in the *hatT::CT* strain, which exhibited a three-log increase in intracellular heme concentration when compared to the *hatT::CT* untreated culture (Figure 11). The intracellular heme concentration was over 40-fold higher in the *hatT::CT* strain treated with heme when compared to the WT strain treated with heme (Figure 11). These data, combined with the heme sensitivity of the *hatT::CT* strain, suggest that the function of HatT is to reduce intracellular heme concentrations to relieve heme toxicity.

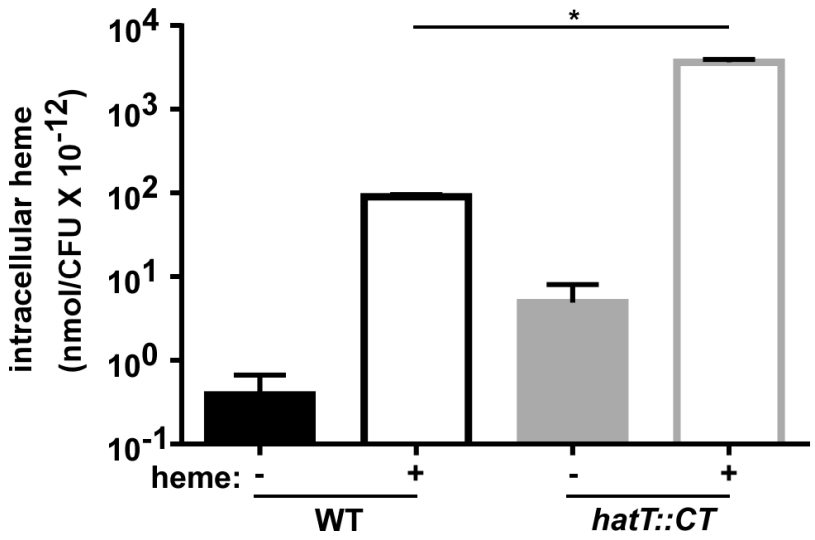

Figure 11 HatT reduces intracellular heme concentrations. WT and *hatT::CT* strains were grown to saturation in either BHIS or BHIS supplemented with 25 μM heme and harvested. Cells were analyzed for their heme content by high resolution MS. Peak areas from extracted-ion chromatograms of heme that accrued above baseline were compared to a standard curve and used to obtain nmoles per CFU. The data are the average of a single experiment performed in biological triplicate with standard deviation. Statistical significance was determined using a multiple comparison one-way ANOVA test with the Tukey correction for multiple comparisons comparing the means of each group to one another. * denotes $p < 0.0001$.

HatT promotes pathogenicity in a mouse model of CDI.

The abundance of heme in the lumen during infection combined with the observed functions of HatR and HatT to sense and reduce heme concentrations, suggest that strains lacking these proteins may have reduced pathogenicity during CDI. To test this, mice were infected with WT, *hatR::CT,* or *hatT::CT* spores and disease was monitored for 4 days. All strains were able to fully colonize the mice as exhibited by ~10^8 colony-forming units (CFU) per gram of stool (Figure 12A). Mice infected with the *hatT::CT* strain lost significantly less weight than the mice infected with the WT or *hatR::CT* strains on days 3 and 4 of the infection (Figure 12B), indicating that the mice infected with the *hatT::CT* strain were partially protected despite similar colonization levels. Furthermore, cecal pathology was significantly reduced in mice infected with the *hatT::CT* strain compared to mice infected with WT or *hatR::CT* strains (Figure 12C). To determine whether the reduced virulence of the *hatT::CT* strain is due to a reduction in toxins TcdA or TcdB, we assessed toxin production in the WT, *hatR::CT*, and *hatT::CT* strains using a cell-rounding cytotoxicity assay. These data revealed toxin levels to be equivalent between all tested strains on day 4 of the infection (Figure 12D), suggesting that the reduced virulence of the *hatT::CT* strain *in vivo* is independent of *C. difficile* toxins. Taken together, these data suggest that the *hatRT* operon senses and detoxifies intracellular heme in *C. difficile* and is required for full pathogenicity during infection.

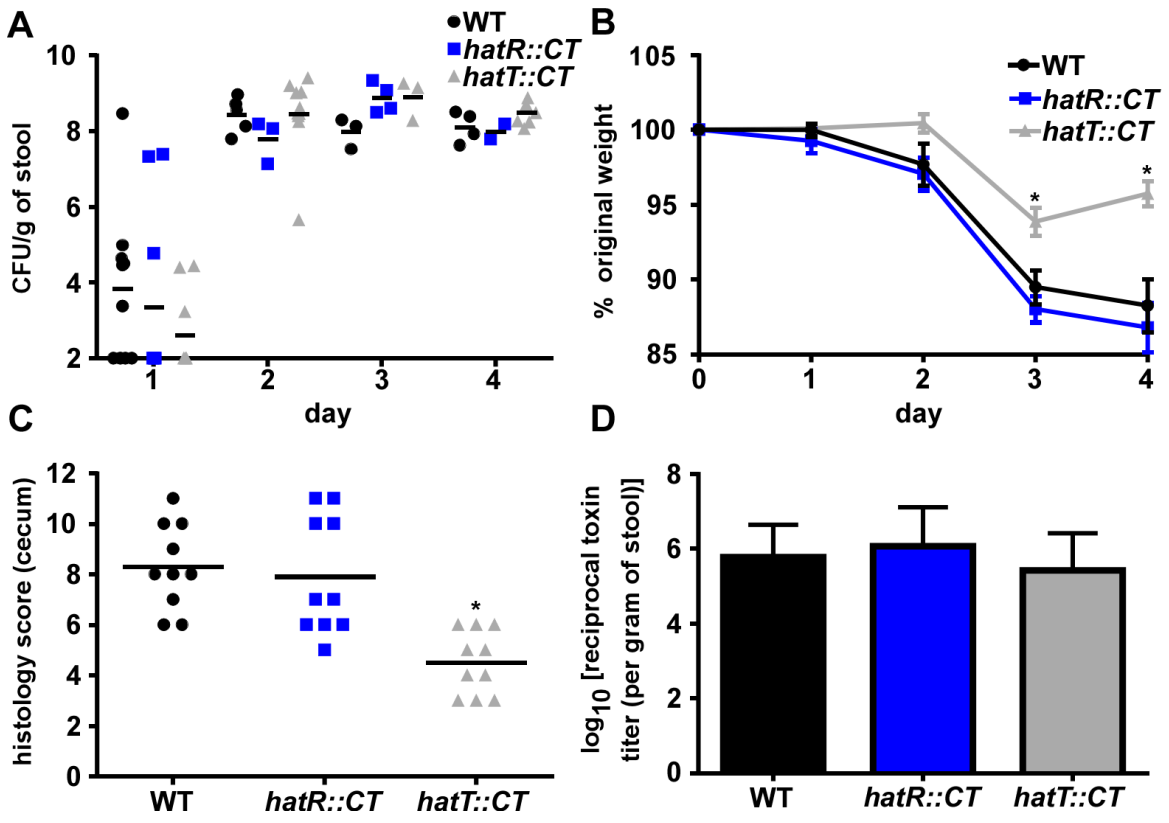

Figure 12 *hatT::CT* displays reduced pathogenicity in a mouse model of CDI. CFU analysis (A) and weights (B) of mice infected with *C. difficile* R20291 WT, *hatR::CT,* and *hatT::CT* strains with standard error on the mean (n = 10/group). (C) Blinded histology scoring of ceca and (D) *C. difficile* toxin titer per gram of feces was measured on day 4 of the infection. All of the data are represented as median or mean with standard error of the mean. Statistical significance was determined using the multiple comparison Kruskal-Wallis test with the Dunn's correction for multiple comparisons comparing the means of each group to one another. * denotes $p < 0.05$.

Discussion

C. difficile infection of the colon causes severe epithelial cell damage, inflammation, and edema, which leads to the hallmarks of *C. difficile*-colitis. Importantly, this damage and subsequent inflammatory response also creates a hostile environment for bacteria within the gut [2, 4, 7, 58]. Highly reactive heme molecules that can be toxic to bacteria are released into the lumen through erythrocyte lysis and necrotic epithelial cell death [8, 11]. Despite the hazard of heme toxicity, *C. difficile* thrives in the colon and survives in the presence of high heme levels. Prior to this work, the mechanism by which *C. difficile* resists heme toxicity were unknown. Herein, I visualized the high abundance of hemoglobin during infection, serving as a proxy for heme, in the murine ceca during CDI. I identified a molecular mechanism encoded by the *hatRT* operon to sense and detoxify heme in *C. difficile.* HatR functions as a transcriptional repressor of the *hatRT* operon and responds to heme concentrations through direct binding of heme. HatR-heme complexes de-repress the *hatRT* operon, leading to the HatT-mediated reduction in intracellular heme concentrations, presumably through efflux. In support of these data, strains with inactivated *hatR* or *hatT* exhibited delayed growth in the presence of heme and the *hatT::CT* strain conferred reduced pathology in a toxin-independent manner in a mouse model of CDI.

While heme sensing and detoxification through efflux is a conserved strategy in multiple Gram-positive organisms, this report is the first to describe an obligate anaerobic pathogen containing such a system [26-28, 59]. TetR-family transcriptional regulators that bind heme have been identified, including HrtR in *Lactococcus lactis,* whereby HrtR regulates heme efflux through a system orthologous to HrtAB [27, 60]. However, HatR shares limited sequence homology (38% amino acid identity) with HrtR. Additionally, the heme binding motifs (single histidine versus two histidines) and the heme-complex disassociation constant (HatR $k_d = 9.2 \pm 1.8$ μM, HrtR $k_d = 0.4$

± 0.2 μM) differ between HatR and HrtR [27]. The significant overexpression of the *hatRT* operon in the presence of heme but not protoporphyrin IX or iron suggests the formation of the heme-HatR complex involves direct binding to the coordinated iron center of heme. The increased heme sensitivity in the *hatR::CT* strain despite the constitutive expression of *hatT,* suggests HatR may also function to reduce heme toxicity through sequestration. The eventual *in vitro* growth observed when the *hatT::CT* strain is exposed to high heme suggests the existence of other mechanisms of heme detoxification in *C. difficile* or the occurrence of suppressor mutations to relieve intracellular heme concentrations through a different transport system. A bioinformatics comparison of HatT with *S. aureus* HrtAB and the dual *S. agalactiae* efflux system PefAB/CD, suggests that these systems arose through convergent evolution as there is little homology between these transporters despite their important role in heme detoxification [28, 59].

The mechanisms of heme toxicity in bacteria are not completely understood. In an anaerobic environment, heme toxicity has been attributed to membrane disruption and DNA damage due to the hydrophobic structure of heme [22-24, 61]. Bilirubin, the terminal metabolite in heme catabolism in mammals is present in high concentrations in the gastrointestinal tract, and destabilizes the membrane of Gram-positive bacteria, suggesting that heme degradation products may also contribute to toxicity [62]. In *C. difficile*, heme enters the intracellular compartment through an unknown mechanism. It is also not known if *C. difficile* utilizes heme as a cofactor or metabolite. Bioinformatic analyses do not reveal heme degradation enzymes of the IsdG or HO enzyme families in *C. difficile* [35, 63, 64]. Additionally, it appears as if *C. difficile* cannot use heme as a sole iron source [39]. In this chapter, we demonstrated that heme accumulates in the cytoplasm of *C. difficile* and is subsequently detoxified through removal by HatT.

Results reported in this chapter demonstrate the importance of heme detoxification in CDI as the *hatT::CT* strain was less pathogenic in a mouse model of infection. The colonization of the WT, *hatR::CT,* and *hatT::CT* strains are at similar levels, supporting a model in which resistance to heme toxicity is important for the end stages of acute infection after serious injury to the intestinal epithelium has occurred. This observation is further supported by the reduction in disease that was only observed on days 3 and 4 following infection in the *hatT::CT* infected mice. The lack of phenotype of in the *hatR::CT* strain suggests that continual expression of *hatT* in the absence of HatR in this strain is sufficient to cause full disease. Surprisingly, there were no differences in bacterial burdens at these days or differences in toxin production despite less overall pathology in the *hatT::CT* infected mice. This suggests *C. difficile* utilizes either additional heme detoxification operons or compensatory mechanisms to relieve intracellular heme stress outside of HatRT and reveals the importance of toxin-independent mechanisms of virulence. Alternatively, as *C. difficile* has been shown to occupy different nutritional niches during infection, and heme is heterogeneously distributed throughout the infected ceca, the heme sensitive strains may be able to maintain WT levels of colonization due to occupying niches of reduced heme concentrations at a cost of pathogenicity [40]. These results provide a molecular insight into how *C. difficile* adapts to the harsh environment of the inflamed gut. Further studies must be performed to elucidate additional mechanisms of protection that *C. difficile* utilizes to survive during infection.

A version of the following section (*Chapter III, CLOSTRIDIOIDES DIFFICILE* SENSES AND HIJACKS HOST HEME FOR INCOPROATION INTO AN OXIDATIVE STRESS DEFENSE SYSTEM) is currently under review at *Cell Host & Microbe* (December 2019).

Knippel, R. J.: de Crécy-Lagard, V. P.: Skaar, E. P. *Clostridioides difficile* Senses and Hijacks Host Heme for Incorporation into an Oxidative Stress Defense System. *Cell Host & Microbe.* Under Review.

CHAPTER III

CLOSTRIDIOIDES DIFFICILE SENSES AND HIJACKS HOST HEME FOR INCORPORATION INTO AN OXIDATIVE STRESS DEFENSE SYSTEM

Introduction

The mechanisms *C. difficile* utilizes to survive in the inflamed colon are largely unknown. In chapter II, I detected a high concentration of hemoglobin from lysed erythrocytes in the gastrointestinal lumen during CDI [65]. Heme, the iron-containing protoporphyrin cofactor of hemoglobin, functions as a redox active cofactor for a number of enzymes, including those involved in respiration and protection against oxidative stress. However, its reactivity causes heme to be toxic, including to *C. difficile*. To cope with heme toxicity, *C. difficile* encodes the HatRT system that senses excess heme and detoxifies the molecule through efflux [65]. HatRT is required for full pathogenicity in a murine model of CDI, underscoring the importance of this process during infection. However, strains inactivated for *hatRT* are not defective in colonization or persistence, suggesting *C. difficile* contains additional mechanisms to survive excess heme encountered during inflammation.

In this chapter, I discovered a *C. difficile* system that senses heme, and utilizes this molecule to provide resistance to oxidative stress. A heme-inducible operon was identified that contains a MarR family transcriptional regulator and a putative membrane protein. I have named these gene products HsmRA for heme sensing membrane protein (R = regulator, A = membrane protein). The total transcriptional response of *C. difficile* to a brief exposure of heme is restricted to the *hsmRA* and *hatRT* operons, and this effect is mediated exclusively by HsmR and HatR. HsmA reduces heme toxicity through sequestration, and heme-bound HsmA provides increased

resistance to compounds that generate oxidative stress, including vancomycin and metronidazole. Lack of HsmA results in reduced colonization persistence in a murine model of relapse infection. Taken together, these results describe a mechanism by which *C. difficile* senses and utilizes heme liberated within the inflamed gastrointestinal tract to provide a protective defense against immune effectors and antibiotic therapy, enabling this organism to thrive at the host-pathogen interface during infection. The conservation of HsmA orthologues across diverse bacterial species suggests that this may be a broadly relevant microbial strategy to survival environmental stress.

Materials and methods

Bacterial strains and plasmids

Bacterial strains used in this study are listed in Table 3. *C. difficile* strains were grown at 37 °C in an anaerobic chamber (85% nitrogen, 10% hydrogen, 5% carbon dioxide, Coy Lab Products) in brain-heart-infusion broth (BD Life Sciences) supplemented with 0.5% yeast extract (BD Life Sciences) and 0.1% cysteine (Sigma-Aldrich) (BHIS) or in *C. difficile* minimal media (CDMM) as described previously [41]. *Escherichia coli* strains were grown in lysogeny broth (LB) or agar (LBA), supplemented with 50 μg/mL kanamycin or 50 μg/mL carbenicillin when necessary. *Bacillus subtilis* strains were grown on LBA or in BHI broth supplemented with 5 μg/mL tetracycline and/or 2.5 μg/mL chloramphenicol. *Staphylococcus aureus* strains were grown on tryptic soy agar (TSA) or in broth (TSB) supplemented with 10 μg/mL chloramphenicol when needed. All antibiotics were purchased from Sigma-Aldrich.

Table 3 Bacterial strains and plasmids used in chapter III.

Bacterial Strain or plasmid	Relevant Feature or Genotype	Reference
Clostridioides difficile R20291		[43]
Clostridioides difficile hsmR::CT	Intron inserted into *hsmR*	This chapter
Clostridioides difficile hsmA::CT	Intron inserted into *hsmA*	This chapter
Bacillus subtilis JH BS2	Carries Tn196	[42]
Escherichia coli DH5α		[44]
Escherichia coli MG1655	RecA+	[45]
Escherichia coli BL21(DE3)		[46]
Staphylococcus aureus RN4220		[66]
Staphylococcus aureus Newman		[67]
Staphylococcus aureus Δ*hrtB*	Deletion of *hrtB*	[68]
Staphylococcus aureus ΔΔ*sod*	Deletion of *sodA* and *sodM*	[69]
Staphylococcus aureus Δ*katA*	Deletion of *katA*	[70]
pJS107	ClosTron plasmid	[42]
pJS107_*hsmR*	ClosTron plasmid with intron targeted to *hsmR*	This chapter
pJS107_*hsmT*	ClosTron plasmid with intron targeted to *hsmA*	This chapter
pJS116	Stable *C. difficile* plasmid	[42]
pJS116_*phsmR-hsmR*	*hsmR::CT* complementation plasmid	This chapter
pJS116_*phsmR-hsmA*	*hsmA::CT* complementation plasmid	This chapter
pOS1	*S. aureus lgt* (constitutive) promoter	[71]
pOS1_*hsmA*	*S. aureus hsmA* expressing plasmid	This chapter
pET15b	Protein expression plasmid	Center for Structural Biology, Vanderbilt University
pET15b_*hsmR*	HsmR expression plasmid	This chapter
pET15b_*hsmR*-H50A	HsmR-H50A expression plasmid	This chapter
pET15b_*hsmR*-K33A	HsmR-K33A expression plasmid	This chapter
pET15b_*hsmR*-K46A	HsmR-H50A expression plasmid	This chapter
pET15b_*hsmR*-Y96A	HsmR-H50A expression plasmid	This chapter
pET15b_*hsmA*	HsmA expression plasmid	This chapter

hsmR::CT and *hsmA::CT* strain generation

Gene inactivations were achieved using the ClosTron system, as described previously [42]. Briefly, gBlocks containing specific modifications for insertion into the genome were generated using the TargeTronics algorithm (http://www.targetrons.com) and synthesized by Integrated DNA Technologies. The gBlocks were cloned into pCR-Blunt vector using the Zero Blunt PCR cloning kit (ThermoFisher Scientific) followed by restriction digest with BsrgI and HindIII (NEB) and ligation (NEB T4 ligase) into pJS107. Plasmids were transformed into the *recA*$^+$ *E. coli* MG1655 through a standard heat shock protocol followed by transformation into *B. subtilis* JH2 using an established method [42]. *B. subtilis* strains containing the pJS107_*hsmR* or pJS107_*hsmA* plasmids were mated with *C. difficile* R20291 overnight at 37 °C by plating and mixing together 100 µL of each strain onto a BHIS plate in the anaerobic chamber. Plates were scraped and transferred into 2 mL of BHIS prior to plating 200 µL onto BHIS plates containing 20 µg/mL thiamphenicol and 50 µg/mL kanamycin ($BHIS_{thiamp20kan50}$). Colonies from these plates were patched onto new $BHIS_{thiamp20kan50}$ and BHIS plates containing 5 µg/mL tetracycline ($BHIS_{tet5}$). Patched colonies that were tetracycline sensitive were patched again onto new $BHIS_{thiamp20kan50}$ and $BHIS_{tet5}$ plates. Colonies that remained tetracycline sensitive were streaked onto BHIS plates containing 20 µg/mL lincomycin ($BHIS_{linc20}$). Inactivation of the *hsmR* or *hsmA* gene was confirmed by performing PCR to identify a 1.5 kbp shift in size using gDNA extracted as previously described on colonies that were lincomycin resistant [42].

Complementation plasmids

Complementation plasmids (Table 3) were created by GenScript using the pJS116 plasmid as a backbone for the synthesized intergenic (271 bp) and full coding region of *hsmR,* and intergenic

region of *hsmR* fused to the full coding region of *hsmA*. *C. difficile* strains were transformed as described above with the removal of the lincomycin selection and were maintained on $BHIS_{thiamp20}$ to ensure plasmid retention.

Protein expression plasmids

Protein expression plasmids for HsmR and HsmA (Table 3) were generated by amplifying *hsmR* flanked by BamHI and NdeI or *hsmA* flanked by XhoI and NedI (Table 4) and cloning into the multiple cloning site of pET-15B after restriction digest. Point mutant generation in pET-15B_*hsmR* was performed with NEB Q5 Site Directed Mutagenesis kit according to the manufacturer's instructions, using the primers listed in Table 4.

Heme, paraquat, and antibiotic toxicity *C. difficile* growth assays

Freshly streaked *C. difficile* colonies were used to inoculate 5 mL of BHIS or $BHIS_{thiamp20}$ and grown for 16 h at 37 °C. Cultures were subcultured 1:50 into fresh BHIS or $BHIS_{thiamp20}$ and grown for 6 h at 37 °C prior to 1:50 inoculation into BHIS, CDMM or $CDMM_{thiamp20}$ containing heme, PPIX, paraquat, vancomycin, or metronidazole at the indicated concentrations. All growth assays were performed in a 96-well plate in 200 µL of media at 37 °C. Optical density at 600 nm (OD_{600}) served as a measurement of growth and was measured every 30 min for the indicated total time in an EpochII microplate reader (BioTek).

Table 4 Oligonucleotides used in chapter III.

Name	Sequence (5' – 3')	Description
qRT_*hsmR*_F	cggtttcaggcataatcagc	qRT-PCR forward primer for *hatR*
qRT_*hsmR*_R	tggcaaattcaaatcctgttg	qRT-PCR reverse primer for *hatR*
qRT_*hsmA*_F	gccaacagccatacttttgaag	qRT-PCR forward primer for *hatR*
qRT_*hsmA*_R	gcccaacttgcatgaaaaag	qRT-PCR reverse primer for *hatR*
qRT_*hatR*_F	ggaaaatagctcaaaggtgtgg	qRT-PCR forward primer for *hatR*
qRT_*hatR*_R	cgccttatcaataaatcccattc	qRT-PCR reverse primer for *hatR*
qRT_*hatT*_F	ttggaaagccctcaactcc	qRT-PCR forward primer for *hatT*
qRT_*hatT*_R	ttttgccatttcagggtctg	qRT-PCR reverse primer for *hatT*
qRT_*rpoB*_F	tgctgttgaaatggttcctg	qRT-PCR housekeeping gene forward primer
qRT_*rpoB*_R	cggttggcatcatcattttc	qRT-PCR housekeeping gene reverse primer
R20291_*hsmR*_F	cggtttcaggcataatcagc	Forward primer to check for intron insertion into *hsmR*
R20291_*hsmR*_R	tgtggcaaattcaaatcctg	Reverse primer to check for intron insertion into *hsmR*
R20291_*hsmA*_F	ttatgctttttcatgcaagttg	Forward primer to check for intron insertion into *hsmA*
R20291_*hsmA*_R	gtaaaagtgtgaaagaaggatgtag	Reverse primer to check for intron insertion into *hsmA*
*hsmR*_NedI_F	ggggcatatggggatgattttattgaaaagtaa	Forward primer for cloning *hsmR* into pET15b
*hsmR*_BamHI_R	ggggatccttatctctcctttacttttgcc	Reverse primer for cloning *hsmR* into pET15b
*hsmR*_H50A_F	aaaactggtagctgctaatcaagaac	Forward primer for H50A point mutation in *hsmR*
*hsmR*_H50A_R	aaaactggtagctgctaatcaagaac	Reverse primer for H50A point mutation in *hsmR*
*hsmR*_K33A_F	gaagaaaattttgcagaaaatggattaac	Forward primer for K33A point mutation in *hsmR*
*hsmR*_K33A_R	gttaatccattttctgcaaaattttcttc	Reverse primer for K33A point mutation in *hsmR*
*hsmR*_K46A_F	gattatagttatagcactggtagctcataatc	Forward primer for K46A point mutation in *hsmR*
*hsmR*_K46A_R	gattatgagctaccagtgctataactataatc	Reverse primer for K46A point mutation in *hsmR*
*hsmR*_Y96A_F	gataaaagaaatacagccgttaagtttac	Forward primer for Y96A point mutation in *hsmR*
*hsmR*_Y96A_R	gtaaacttaacggctgtatttcttttatc	Reverse primer for Y96A point mutation in *hsmR*
*hsmA*_NedI_F	gcccatatgaattataaattaatacttgc	Forward primer for cloning *hatR* into pET15b and pOS1
*hsmA*_BamHI_R	ctatccaaccataccggatcc	Reverse primer for cloning *hatR* into pET15b and pOS1

qRT-PCR

C. difficile were grown anaerobically in triplicate in CDMM at 37 °C to an OD_{600} of 0.3 abs. Hemin (Sigma) was solubilized in 0.1 M NaOH and added to 50 μM or the indicated concentration. After 5 or 30 min, a 1:1 solution of acetone:ethanol was added to an equal volume of the culture. For activation analysis, *C. difficile* were grown in CDMM containing the indicated concentrations of NaOH, protoporphyrin-IX, iron (II) sulfate, or heme to an OD_{600} of 0.3 abs prior to addition of acetone:ethanol. Samples were stored at -80 °C until used for RNA extraction. Samples were thawed on ice, pelleted, and resuspended in 750 μL of LETS buffer (1 M LiCl, 0.5 M EDTA, 1 M Tris pH 7.4). Cells were transferred to tubes containing lysing matrix B beads (MP Biomedicals) and lysed by a FastPrep-24 (MP Biomedicals) bead beater for 45 s at 6 m/s. Lysed samples were heated for 5 min at 55 °C and pelleted by centrifugation for 10 min. The supernatant was transferred to a fresh tube and 1 mL TRIzol (Thermo Scientific) was added. Chloroform (200 μL) was added to each sample and vortexed prior to separation of the aqueous and organic layers by centrifugation for 15 min. The aqueous (upper) layer was transferred to a fresh tube and the RNA was precipitated through the addition of 1 mL isopropyl alcohol. Samples were incubated for 10 min and RNA was pelleted by centrifugation for 10 min. Supernatant was removed and the RNA pellet was washed with 200 μL of 70% ethanol. Samples were air dried for 1 min, then resuspended in 100 μL RNase free water. DNA contamination was removed through the addition of 8 μL RQ1 DNase, 12 μL 10x RQ1 buffer, and 2 μL RNase inhibitor (Promega) to the purified RNA. Samples were DNase treated for 2 h and purified using the RNeasy miniprep RNA cleanup kit (Qiagen). RNA concentration was determined using the Synergy 2 with Gen 5 software (BioTek) and 2 μg was reverse transcribed by M-MLV reverse transcriptase (Fisher Scientific) in the presence of RNase inhibitor (Promega) and random hexamers (Promega). Reactions lacking

the reverse transcriptase were used to control for DNA contamination. Newly created cDNA was diluted 1:100 and was used in qRT-PCR using iQ SYBR green supermix (BIO-RAD) utilizing the primer pairs in Supplementary Table 2. Amplification was achieved using a 3-step melt cure program on a CFX96 qPCR cycler (BIO-RAD). Transcript abundance was calculated using the ΔΔCT method normalized by the *rpoB* gene.

HsmR multiple sequence alignment

HsmR multiple sequence alignment was performed using Clustal Omega (http://www.clustal.org/omega [accessed August 2019]) using *Clostridioides difficile* R20291 (NCBI accession no. CBE02826.1), *Clostridium perfringens* ATCC 13124 (NCBI accession no. ABG82197.1), *Clostridium botulinum* A str. Hall (NCBI accession no. YP_001388019.1) and *Clostridium novyi* (NCBI accession no. WP_011722139.1).

HsmA taxonomic distribution and physical clustering

HsmA **(**CDR20291_0781) is part of the Interprofamily IPR023813 that contains 750 proteins with only ~ 100 from *Clostridioides* or *Clostridium* derivatives. To visualize the genetic spread of the family the corresponding protein sequence was used as input to search all representative and reference genomes using the internal BlastP search tool in the Patric database version 3.5.43 [72]. One hundred eighty one proteins with alignment scores > 55 were extracted and were confirmed for membership to the IPR023813 family. Forty-two of these were found in organisms with complete genome sequences. This group of 42 bacteria was merged to the group of 120 reference genomes present in the Patric database and these 157 genomes (5 were found in the two groups) were used to build a species phylogeny using the internal Patric CodonTree tool. The output file

in newick format was used as input in Itol v4 with added HsmA homolog presence/absence data [73]. Physical clustering analysis was performed using the gene neighborhood tool of PubSEED and can be visualized in the CD0851 SubSystem (http://pubseed.theseed.org//SubsysEditor.cgi?page=ShowSpreadsheet&subsystem=CD0851) [74].

Protein expression and purification

E. coli BL21 (DE3) pREL containing the pET-15b_*hsmR* plasmids were grown overnight in 5 mL of LB_{carb50} at 37 °C. Cells were subcultured into Terrific broth (ThermoFisher Scientific) containing 50 µg/mL carbenicillin and grown to the mid-logarithmic phase of growth (0.5 abs measured at 600 nm) at 37 °C prior the addition of 1 mM isopropyl-1-thiol-D-galactopyranoside (IPTG). Growth was continued at 16 °C for 16 h. Cells were harvested by centrifugation (6000 x *g* for 10 min) and resuspended in 1 X PBS. Cells were lysed by passage through an EmulsiFlex homogenizer (Avestin) three times at 20,000 lb/in^2. The insoluble debris was removed by centrifugation at 40,000 x *g* for 1 h and the supernatant was filtered using a 0.22-µM-pore sizer filter. Filtered lysate was added to HisPur cobalt resin (ThermoFisher Scientific) and allowed to bind at 4 °C for 30 min prior to transfer to a gravity column. The column was washed with four column volumes of wash buffer (100 mM HEPES, 500 mM NaCl, pH 7.8) three times followed by 2 column volumes of elution buffer (100 mM HEPES, 500 mM NaCl, 200 mM imidazole, pH 7.8) twice. The hexahistidine tag was cleaved using the Thrombin Cleancleave kit (Sigma-Aldrich) by following the manufacturer's instructions. After cleavage, buffer was exchanged utilizing overnight dialysis at 4 °C in 4 L of wash buffer (100 mM HEPES, 500 mM NaCl, pH 7.8).

Solubilization of membrane fractions

E. coli BL21 (DE3) pREL containing the pET-15b_*hsmA* plasmids were grown overnight in 5 mL of LB_{carb50} at 37 °C. Cells were subcultured into 25 mLs of fresh LB_{carb50} and grown to the mid-logarithmic phase of growth (0.5 abs measured at 600 nm) at 37 °C prior to the addition of 1 mM IPTG and/or 10 μM heme. Growth was continued at 16 °C for 16 h. Cells were harvested by centrifugation (6000 x *g* for 10 min), photographed using a dual 12-megapixel camera (Apple) and resuspended in 1 X PBS. Membranes were solubilized for 1 h at 4 °C by adding octyl-β-glucoside to a final concentration of 1.5% with gentle rocking. Insoluble fraction was pelleted by centrifugation (20,000 x g for 3 min) and soluble membrane fraction was removed.

Absorption spectroscopy

Heme binding by HsmR was determined by measuring the absorption spectrum of increasing amounts of hemin (0 – 30 μM) after addition to a cuvette containing 10 μM recombinant HsmR in 1 mL of Tris-buffered saline (TBS) and a reference standard containing 1 mL TBS on a Varian Cary 50BIO. Samples were mixed and allowed to incubate at room temperature in the dark for 5 min prior to collecting the spectrum between 300 – 800 nm with 10 nm increments. Binding ratio of heme to HsmR was determined by plotting the change in absorbance at 413 nm between the reference standard and the HsmR sample. A curve fit and ratio was obtained by performing the one-site binding model non-linear regression function on Graph Pad Prism 8.2. Heme binding by HsmA was performed on 1 mL of solubilized membrane fractions isolated as described above collecting the spectrum between 300 – 800 nm with 10 nm increments.

Heme and oxidative stress toxicity *S. aureus* growth assays

Freshly streaked *S. aureus* colonies were used to inoculate 5 mL of TSB or TSB_{cm10} in 15 mL round-bottom polypropylene tubes with aeration lids and grown for 16 h at 37 °C at a 45° angle in an Innova 44 incubator shaking at 180 rpm. Cultures were subcultured 1:50 into fresh TSB or TSB_{cm10} and grown for 6 h at 37 °C prior to 1:50 inoculation into TSB or TSB_{cm10} containing heme, paraquat, or hydrogen peroxide at the indicated concentrations. All growth assays were performed in a 96-well plate in 200 µL of media shaking linearly at 567 cpm (3 mm) at 37 °C. Optical density at 600 nm (OD_{600}) served as measurement of growth and was measured every 30 min for the indicated total time in an EpochII microplate reader (BioTek).

RNA-sequencing analysis

RNA was isolated and purified as described above. RNA sequencing was performed by the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core using the Illumina HiSeq 3000 platform (Illumina). The integrity and concentration of total RNA were determined using an Agilent 2100 Bioanalyzer system in combination with an RNA 6000 Nano kit (Agilent). rRNA was depleted using the Ribo-Zero rRNA removal kit (for bacteria) (Epicentre) and paired-end cDNA libraries were prepared with a TruSeq RNA library prep kit v2 (Illumina). Data analysis for sequencing experiments was performed on the CLC Genomics workbench (version 11.0.1; Qiagen) using the reference *C. difficile* R20291 genome. Prior to analysis, rRNA reads were removed in order to account for variations in rRNA depletion procedure among samples. Standard settings were used for adapter and quality trimming, as well as transcriptome sequencing (RNA-seq) analysis. Expression values were calculated as RPKM (reads per kilobase per million mapped

reads) [47], and a lower cutoff of 5 RPKM was introduced for subsequent analysis. Raw sequence data are deposited on the NCBI Sequence Read Archive (accession code: PRJNA576216).

Relapse mouse model of CDI

All animal experiments under protocol M1700053 were reviewed and approved by the Institutional Animal Care and Use Committee of Vanderbilt University. Procedures were performed according to the institutional policies, Animal Welfare Act, NIH guidelines, and American Veterinary Medical Association guidelines on euthanasia. Adult (6 – 8 week old) age-matched male C57Bl/6 (Jackson Laboratories) mice were housed in groups of five and maintained at Vanderbilt University Medical Center Animal Facilities. Mice were subjected to a previously described model of CDI [75]. Briefly, mice were treated with 0.5 mg/mL cefoperazone in their drinking water for 5 days. Mice were given a 2 day recovery period prior to administration of 10^5 spores of WT, *hsmR::CT,* or *hsmA::CT C. difficile* strains in PBS via oral gavage. Prior to infection, mice were confirmed to be *C. difficile* negative. After infection, mice were monitored for signs of disease, including diarrhea and weight loss. At day 4 post infection, mice were treated with 0.2 mg/mL vancomycin in their drinking water for 5 days. Mice were monitored for relapse after removal of vancomycin. On the final day of infection and necropsy, cecal contents were harvested. Mice that displayed severe disease or weight loss greater than 20% were humanely euthanized.

Bacterial burden determination

C. difficile CFUs were quantified daily from fecal or cecal samples. Samples were diluted and homogenized in PBS and serial plated onto taurocholate cycloserine cefoxitin fructose agar (TCCFA) for enumeration as CFU per gram of feces.

Reactive oxygen species measurements

Freshly streaked *C. difficile* colonies were used to inoculate 5 mL of BHIS grown for 16 h at 37 °C. Cultures were subcultured 1:50 into fresh BHIS and grown for 6 h at 37 °C prior to the addition of heme and dihydrorhodamine 123 (Invitrogen) for 30 min at the indicated concentrations. Two hundred μL of culture was transferred to a 96-well plate and sealed with a Breathe-Easy gas permeable membrane (Diversified Biotech). Sealed plates were removed from the chamber and OD_{600} and fluorescence (excitation = 507 nm; emission = 529 nm) was measured at 10 min intervals on a Cytation 5 (BioTek) shaking in a double orbital at 567 cpm (3 mm) at 37 °C in atmospheric oxygen.

S. aureus pOS1 and pOS1_*hsmA* were diluted 1:50 into 200 μL TSB or TSB_{cm10} containing paraquat and dihydrorhodamine 123 (Invitrogen) at the indicated concentrations. OD_{600} and fluorescence (excitation = 507 nm; emission = 529 nm) were measured at 10 min intervals on a Cytation 5 (BioTek) shaking linearly at 567 cpm (3 mm) at 37 °C. The data displayed are background corrected for the wells with all components except cells and normalized to OD_{600}.

Results

Transcription of the *hsmRA* operon occurs rapidly in response to heme.

Analysis of RNA-sequencing of transcripts from *C. difficile* R20291 exposed to a sublethal concentration of heme revealed two uncharacterized genes as being highly responsive to heme exposure [65]. More specifically, two genes in a candidate operon encoding for a MarR family transcriptional regulator (CDR20291_0782) and a putative membrane protein with homology to a

cytochrome b561 (CDR20291_0781) exhibited the most significant changes in transcript abundance in this data set (Figure 13A). To confirm CDR20291_0782 and CDR20291_0781 are heme responsive, quantitative reverse transcription PCR (qRT-PCR) was performed utilizing RNA harvested at the early exponential phase of growth (OD_{600} = 0.3) from cultures grown in equimolar concentrations of sodium hydroxide (NaOH, vehicle), protoporphyrin IX (PPIX, porphyrin ring without iron), iron (II) sulfate, or heme. Transcription of each gene was minimally altered in the samples treated with NaOH or iron (II) sulfate, in contrast to a 1-log increase in transcript abundance following protoporphyrin IX exposure, and 2 - 2.5-log increase in transcript abundance following heme exposure (Figure 13B). To examine the responsiveness of the *hsmRA* and *hatRT* operons to heme, qRT-PCR was performed on RNA harvested from cultures grown to an early exponential phase of growth (OD_{600} = 0.3) and exposed to a range of low concentrations of heme (0.25 – 1 μM) for 5 min. There was no significant difference in the transcription of these four genes at 0.25 and 0.5 μM heme (Figure 13C). At 0.75 and 1 μM heme, transcription of CDR20291_0782 and CDR20291_0781 was 1 – 2-log higher than an untreated control. This transcriptional response was more rapid and intense than that of the previous characterized heme efflux system *hatRT,* [65] as this operon displayed minimal transcriptional change at 0.75 μM and a 1-log transcriptional increase at 1 μM heme compared to an untreated control (Figure 13C). Based on these data, as well as data described below, I have named CDR20291_0782 heme sensing membrane protein regulator (*hsmR*) and CDR20291_0781 heme sensing membrane protein (*hsmA*).

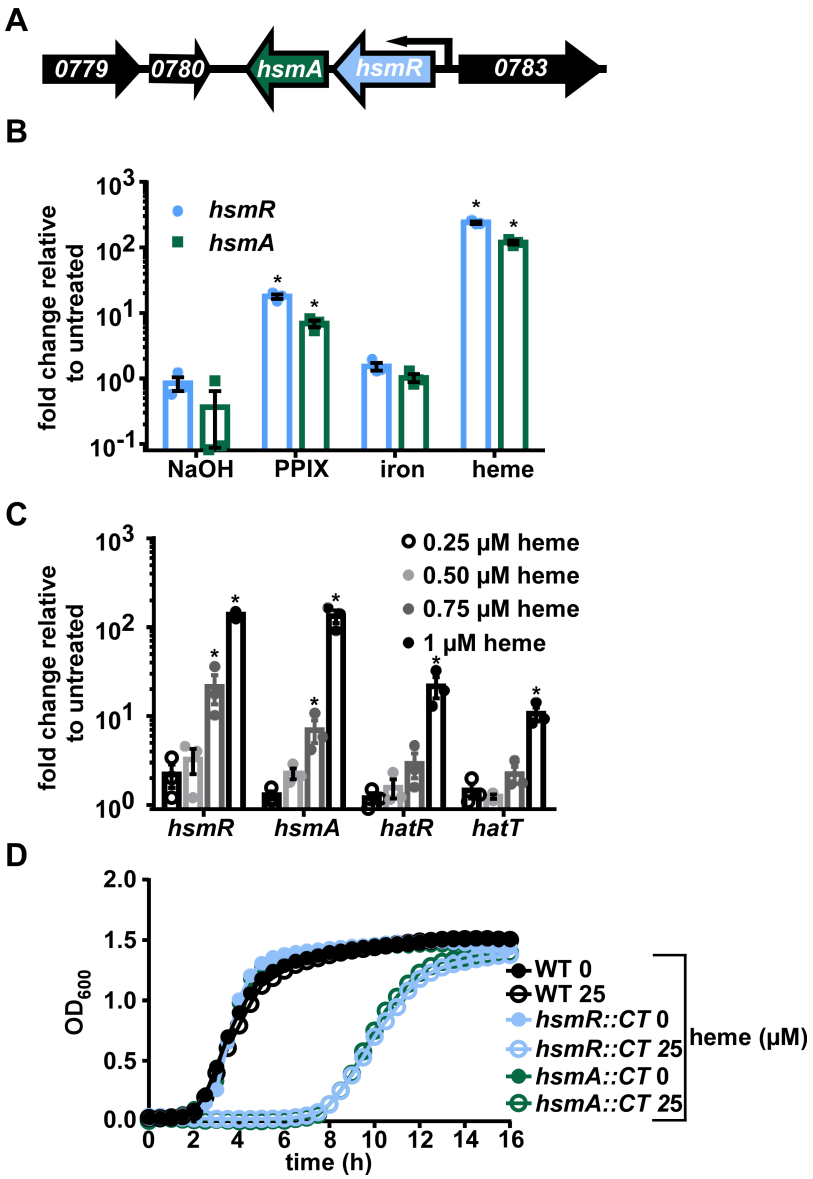

Figure 13 *hsmRA* transcriptionally responds to and detoxifies heme.
(A) Schematic of the *hsmRA* operon. (B) *hsmR* and *hsmA* transcription determined by qRT-PCR. cDNA was reverse transcribed from RNA harvested from *C. difficile* R20291 grown in the presence of sodium hydroxide (NaOH, 500 µM), protoporphyrin IX (PPIX, 50 µM), iron sulfate (50 µM) or heme (50 µM). (C) *hsmR, hsmA, hatR,* and *hatT* transcription determined by qRT-PCR of cDNA reverse transcribed from RNA harvested from *C. difficile* grown in the presence of a low concentration range of heme (0.25 – 1 µM). Transcription is graphed as the fold change relative to an untreated control. The data are a representative of three independent experiments each in biological triplicate with standard deviation shown. Statistical significance was determined using the multiple comparison one-way ANOVA test comparing the means of each group to one another * denotes $p < 0.05$. (D) Growth of WT, *hsmR::CT,* and *hsmA::CT* strains in the presence or absence of heme (25 µM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

HsmR and HsmA reduce heme toxicity.

To investigate the contribution of the *hsmRA* operon to heme detoxification, two strains of *C. difficile* were generated that are inactivated for *hsmR* (*hsmR::CT*) or *hsmA* (*hsmA::*CT). The lack of either *hsmR* or *hsmA* renders the bacteria sensitive to heme toxicity, as growth over time in the presence of 25 μM heme is delayed in the mutant strains compared to wild-type (WT, Figure 13D). The growth of *hsmR::CT* and *hsmA::CT* returned to WT levels by expressing the relevant gene *in trans* under the control of their native promoters (Figure 14). Taken together, these data suggest that HsmR and HsmA coordinate to function as an acutely sensitive mechanism of heme sensing and detoxification in *C. difficile.*

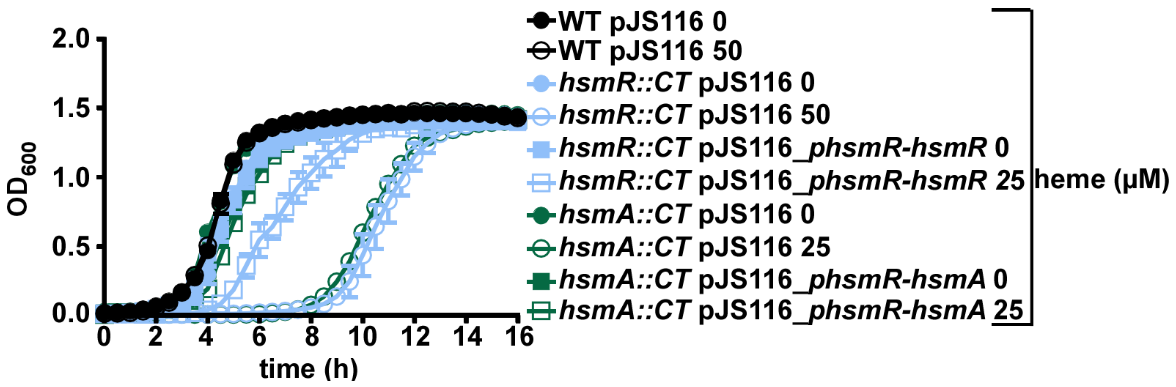

Figure 14 Heme sensitivity in WT, *hsmR::CT, hsmA::CT*, and complemented strains. Growth of *C. difficile* WT pJS116 (empty vector), *hsmR::CT* pJS116, *hsmR::CT* pJS116_*phsmR-hsmR*, *hsmA::CT* pJS116, and *hsmA::CT* pJS116_*phsmR-hsmA* strains in CDMM in the presence or absence of heme (25 µM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

HsmR and HsmA coordinate to reduce heme toxicity through sequestration.

Orthologs of *hsmR* were identified utilizing the SEED database in other clostridial species that are pathogenic to vertebrates and therefore may experience heme stress during infection (Figure 15A) [76]. Considering the responsiveness of *hsmRA* to both heme and PPIX, and the candidate assignment of HsmR as a transcriptional regulator, I examined the ability of HsmR to bind heme. Recombinant HsmR (10 μM) was incubated with heme (1 – 30 μM), resulting in the appearance of a Soret peak at 413 nm (Figure 15B), indicative of HsmR-heme complex formation [57]. Differential absorption spectroscopy at 413 nm over a range of heme concentrations was used to determine that HsmR binds heme at a 1:1 ratio using a single site binding model (k_d = 6.6 ± 1.1 μM; Fig. 2b insert). The conserved histidine residue at position 50 was identified as a potential axial ligand to bind heme (Figure 15A red box). Purified recombinant HsmR containing the substitution of histidine 50 to alanine (H50A) exhibited a reduced ability to bind heme (Figure 15C and Figure 16). Alanine substitutions of other conserved residues with the potential to bind heme (K33A, K46A, and Y86A) did not significantly alter heme binding (Figure 17).

Based on the homology between HsmA and heme-containing cytochromes, I investigated the ability of HsmA to bind heme. Attempts to purify HsmA were unsuccessful; therefore, we developed a whole cell assay to measure HsmA heme binding. HsmA expression was induced by IPTG in an expression strain of *E. coli* (*E. coli* pET15b_*hsmA*), and this strain was treated with excess heme. The presence of excess heme in HsmA expressing strains resulted in a dark red cell pellet, indicative of bound heme, and this color change is not observed in *E. coli* lacking the HsmA expression vector upon heme exposure, or in the un-induced *E. coli* pET15b_*hsmA* strain (Figure 15D). Absorption spectroscopy of the solubilized membranes from these cells resulted in a unique

peak at ~405 nm that was only present in the *E. coli* pET15b_*hsmA* strain treated with both heme and IPTG (Figure 15E). These data suggest that HsmR and HsmA bind heme.

The ability of HsmA to bind heme could reduce toxic free heme concentrations through sequestration, as has been observed in other heme detoxification systems [20]. To test this, a *Staphylococcus aureus* strain lacking the *hrtB* (Δ*hrtB*) heme efflux pump was transformed with a plasmid containing a constitutively expressed *hsmA* (Δ*hrtB* pOS1_*hsmA*). *S. aureus* Δ*hrtB* was utilized for these experiments due to this strain's high sensitivity to heme [22, 25]. The *S. aureus* Δ*hrtB* pOS1_*hsmA* displayed increased resistance to heme toxicity when compared to the Δ*hrtB* strain harboring empty vector (Figure 15F). Together these data suggest that HsmR and HsmA coordinate to reduce heme toxicity through direct binding to HsmA.

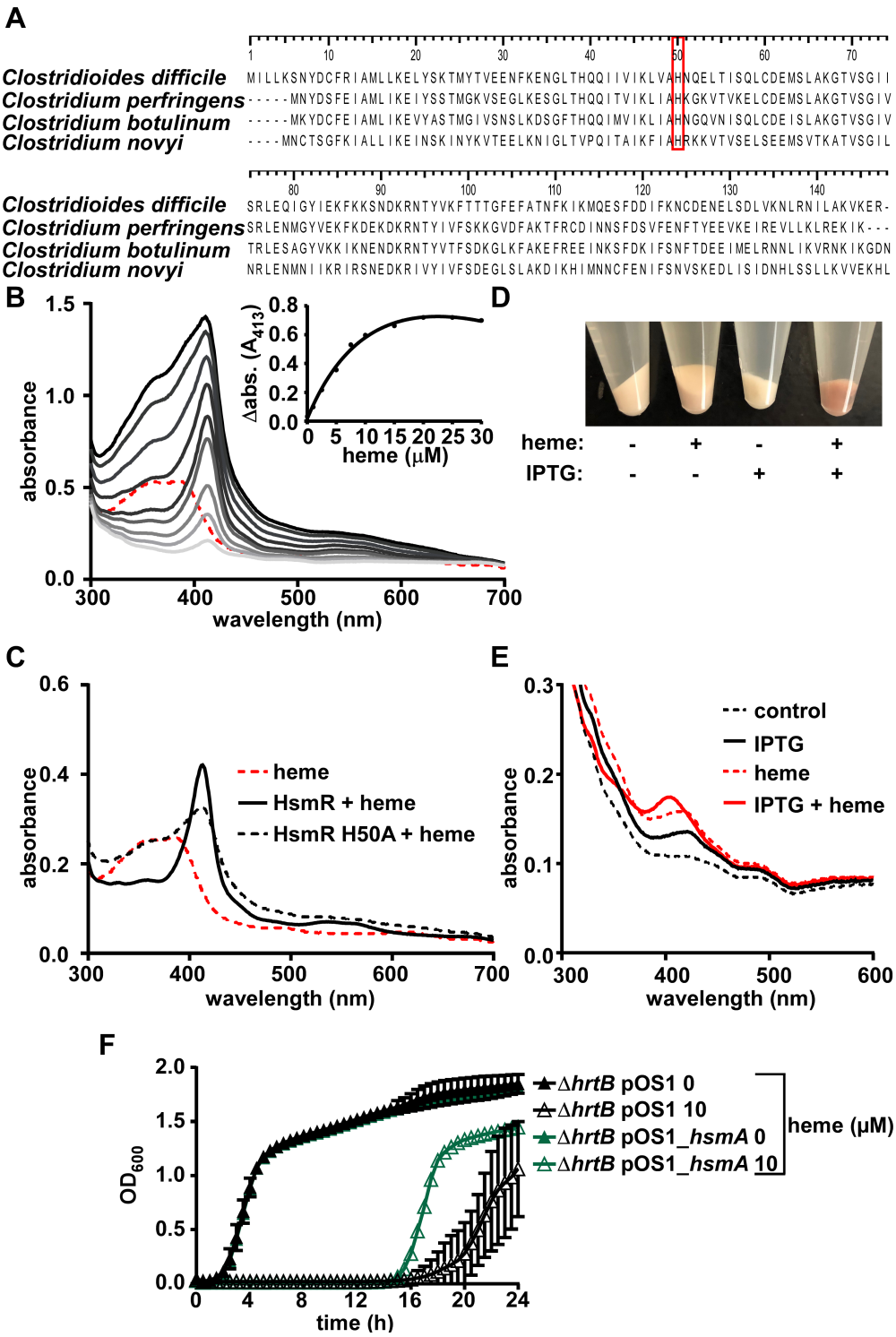

Figure 15 HsmR binds and senses heme and HsmA reduces heme toxicity through sequestration.

(A) Alignment of HsmR homologues in other pathogenic clostridial species. Red box denotes conserved histidine residue. (B) Absorption spectra of heme binding to recombinant HsmR. Increasing concentrations of heme (1 to 30 μM) were added to 10 μM protein. The spectrum corresponding to 10 μM heme is shown as a dashed red line. HsmR with increasing concentrations

of heme are shown as gray lines. The inset displays change in absorbance at 413 nm for HsmR bound to heme minus the corresponding heme alone peak. (C) Absorption spectra of 10 μM heme binding to HsmR or HsmR H50A. (D) *E. coli* pET15b_*hsmA* cell pellets in the presence or absence of heme (10 μM) and IPTG (1 mM). (E) Absorption spectra of solubilized membrane fractions of the cell pellets from D. (F) Growth of *S. aureus* Δ*hrtB* pOS1 and pOS1_*hsmA* strains in the presence or absence of heme (10 μM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

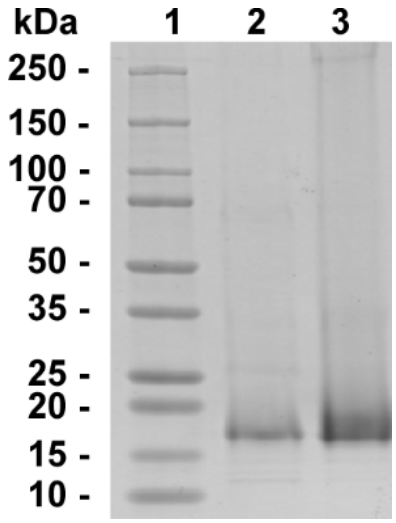

Figure 16 Purified recombinant HsmR and HsmR H50A.
Coomassie stained SDS-PAGE of purified recombinant HsmR and HsmR H50A. 1 = protein ladder. 2 = HsmR (17 kDa). 3 = HsmR H50A (17 kDa).

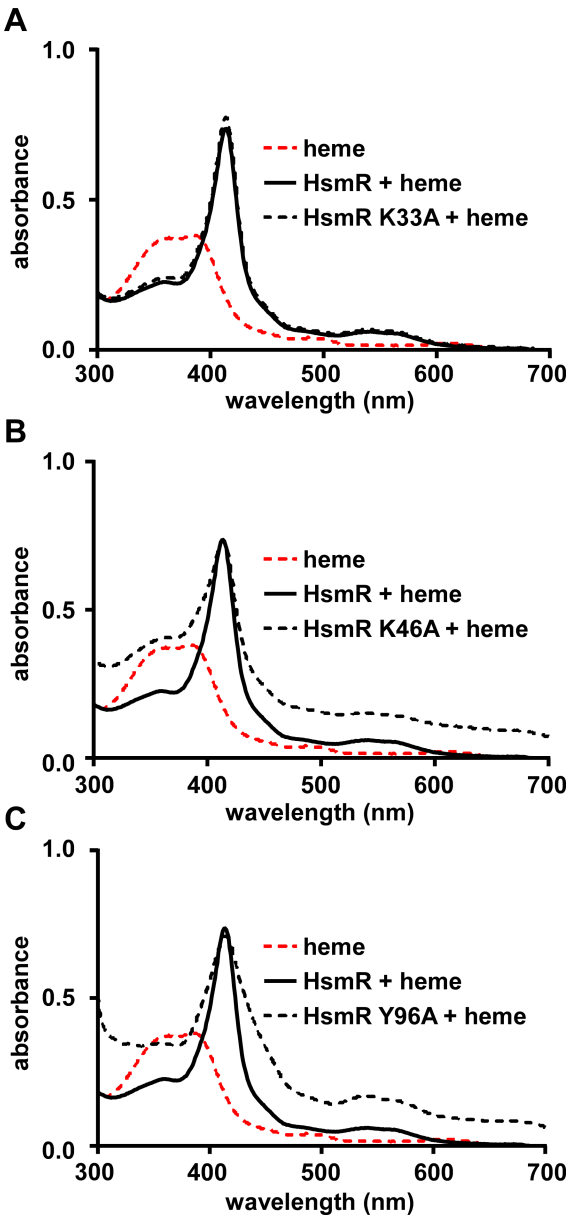

Figure 17 Conserved histidine residues within HsmR that are not required for heme binding. Absorption spectra of 10 μM heme binding to 10 μM HsmR and HsmR K33A (A), HsmR K46A (B), and HsmR Y86A (C).

HsmR is a transcriptional activator of the *hsmRA* operon.

Members of the MarR family of transcriptional regulators can function as repressors or activators of genes [77]. I investigated the regulation of the *hsmRA* operon by HsmR utilizing qRT-PCR of WT and *hsmR::CT* grown to early exponential phase (OD_{600} = 0.3) and exposed to 50 μM heme for 30 min. Heme-treated WT exhibit a 1 – 1.5-log activation of *hsmRA*, and this is entirely dependent on a functional HsmR, as *hsmRA* transcription was unaffected by heme in *hsmR::CT* (Figure 18A). These data suggest HsmR acts as an activator of the *hsmRA* operon.

To identify additional genes regulated by HsmR, I performed an RNA-sequencing experiment comparing the total relative mRNA transcript abundance of WT and *hsmR::CT* grown to early exponential phase (OD_{600} = 0.3) and exposed to 50 μM heme for 30 min. In WT *C. difficile, hsmR, hsmA, hatR,* and *hatT* were the only four genes that displayed a significant induction above $\log_2$ of 2 following exposure to heme (Figure 18B, Appendix B Table 1). The comparison between heme-exposed *hsmR::CT* and the untreated *hsmR::CT* control replicated the observation that HsmR acts as an activator of the *hsmRA* operon. Upregulation of *hsmR* and *hsmA* transcripts were not observed in the *hsmR::CT* strain in response to heme (Figure 18C, Appendix B Table 2), whereas *hatR* and *hatT* retained significant transcriptional induction in this strain. Additional comparisons between untreated and heme treated WT and *hsmR::CT* samples further revealed the absence of significant heme induced transcriptional changes outside of the *hatRT* and *hsmRA* operons (Figure 19, Appendix B Tables 3 and 4). Taken together, these data demonstrate that the heme responsive regulon of HsmR is solely comprised of *hsmRA*. Moreover, these studies show that the transcriptional response of *C. difficile* to heme is limited to the *hsmRA* and *hatRT* operons.

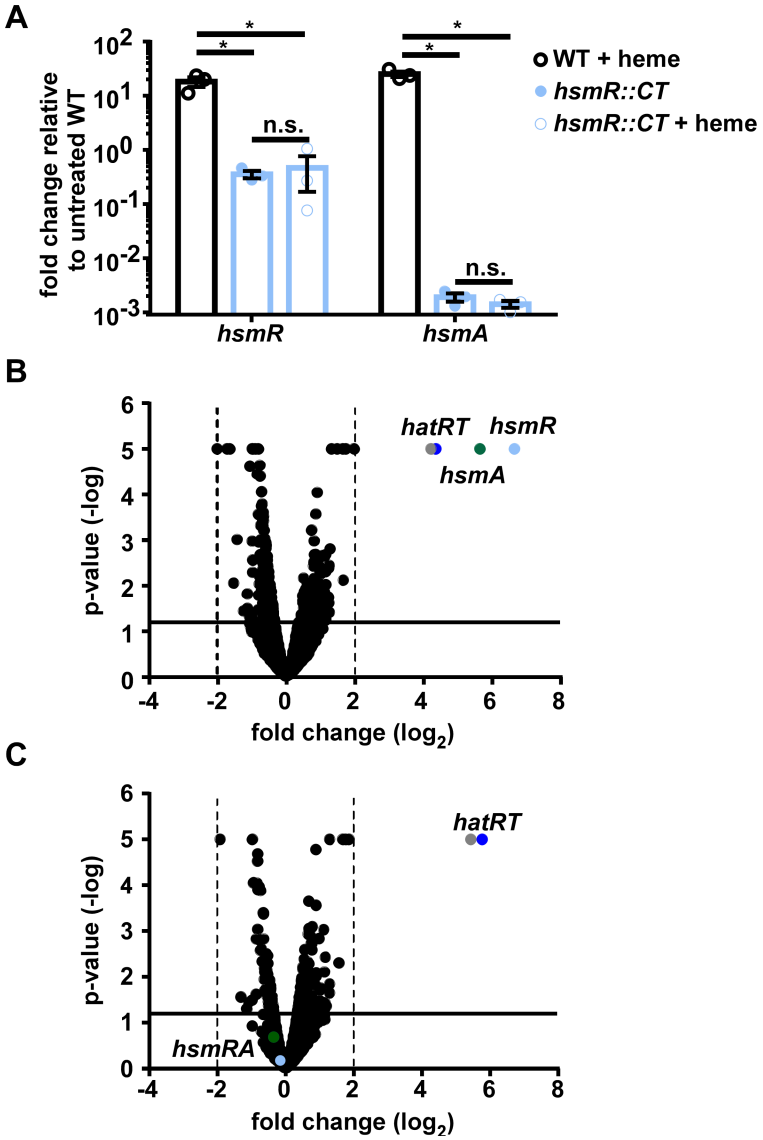

Figure 18 HsmR acts as an activator of the *hsmRA* operon.
(A) *hsmR* and *hsmA* transcription determined by qRT-PCR. cDNA was reverse transcribed from RNA harvested WT or *hsmR::CT* grown to early exponential phase (0.3 abs) and exposed to heme (50 µM) for 30 min. (B) RNA-sequencing analysis comparing RNA from heme treated (25 µM for 30 min) WT to an untreated WT control. (C) RNA-sequencing analysis comparing RNA from heme treated (25 µM for 30 min) *hsmR::CT* to an untreated *hsmR::CT* control. Dashed lines represent genes of fold change > 2. Samples with *p*-value > 1 x 10^{-5} are represented as 5 on the graph. Solid black line denotes $p < 0.05$. Statistical significance was determined using the multiple comparison two-way ANOVA test with the Sidak correction for multiple comparisons comparing the means of each group to one another The data are a representative from three independent experiments each in biological triplicate with standard error of the mean. * denotes *p*-value < 0.05, n.s. denotes not significant.

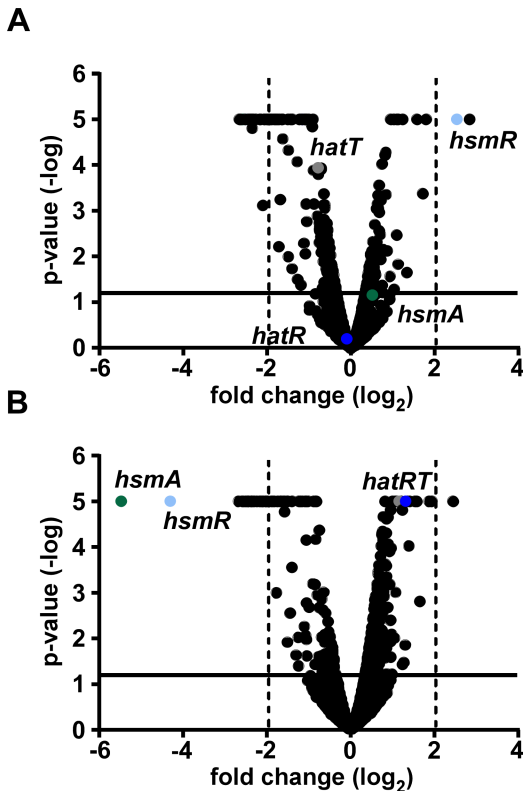

Figure 19 The regulon of HsmR is limited to the *hsmRA* operon.
(A) RNA-sequencing analysis comparing RNA from heme treated (25 µM for 30 min) *hsmR::CT* to heme treated (25 µM for 30 min) WT. (B) RNA-sequencing analysis comparing RNA from untreated *hsmR::CT* to untreated WT. Dashed lines represent genes of fold change > $Log_2(2)$. Samples with *p*-value > 1 x 10^{-5} are represented as 5 on the graph. Solid black line denotes $p < 0.05$. Statistical significance was determined using the multiple comparison one-way ANOVA test with the Tukey correction for multiple comparisons comparing the means of each group to one another. The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

HsmA employs exogenous heme to confer resistance to oxidative stress and antibiotics.

HsmA contains homology to the cytochrome b561 family, but *C. difficile* does not respire [78]. Therefore, I investigated other functions of cytochromes to determine the functional role of this protein. One function of cytochromes is to detoxify redox molecules by shuttling electrons through bound heme cofactors [79]. I investigated whether HsmA may function to diminish damage caused by redox active molecules [80]. To test this, WT, *hsmR::CT,* and *hsmA::CT* cultures were treated with the ROS-generating molecule paraquat in the presence or absence of heme [81]. No toxicity was observed when *C. difficile* was treated with paraquat alone (2 mM; Figure 20), whereas a delay in growth occurred in heme and paraquat treated *hsmR::CT* and *hsmA::CT,* but not WT (Figure 20).

Considering *C. difficile* is an obligate anaerobe, and HsmA protects this organism against superoxide stress, I sought to investigate the contribution of HsmA to survival in the presence of atmospheric oxygen. WT and *hsmA::CT* were grown for 6 hours prior to treatment with heme (25 μM) for 30 min. Samples were exposed to atmospheric oxygen, and the heme treated WT samples displayed a 2.5-log decrease in cellular ROS compared to the untreated WT strain (Figure 21A). By contrast, heme-treated *hsmA::CT* displayed a 2.5 log increase in cellular ROS compared to the untreated *hsmA::CT* strain (Figure 21A). There were no statistically significant differences in cellular ROS between the untreated WT and *hsmA::CT* strains (Figure 21A). To decouple the ability of HsmA to reduce oxidative stress from other oxidative stress reducing proteins in *C. difficile,* a *S. aureus* strain lacking both genes encoding for superoxide dismutase enzymes (ΔΔ*sod*) was transformed with a plasmid containing a constitutively expressed *hsmA* (ΔΔ*sod* pOS1_*hsmA*) or with an empty vector control plasmid (ΔΔ*sod* pOS1). *S. aureus* ΔΔ*sod* is acutely sensitive to superoxide stress [82]. The ΔΔ*sod* pOS1_*hsmA* strain displayed a recovery in aerobic growth over

time in 2 mM paraquat when compared to the ΔΔ*sod* pOS1 strain (Figure 21B). Measuring the cellular ROS generated at 6 h displayed a 3-fold decrease in ΔΔ*sod* pOS1_*hsmA* strain compared to the ΔΔ*sod* pOS1 strain (Figure 21C). To determine if the protective effect is specific to superoxide, a *S. aureus* strain lacking catalase (Δ*katA*) was transformed with a plasmid containing a constitutively expressed *hsmA* (Δ*katA* pOS1_*hsmA*) or empty vector control (Δ*katA* pOS1) [83]. The expression of HsmA did not confer resistance to hydrogen peroxide (Figure 22). Taken together, these data suggest HsmA, in addition to detoxifying heme, provides protection against atmospheric and chemically generated superoxide stress.

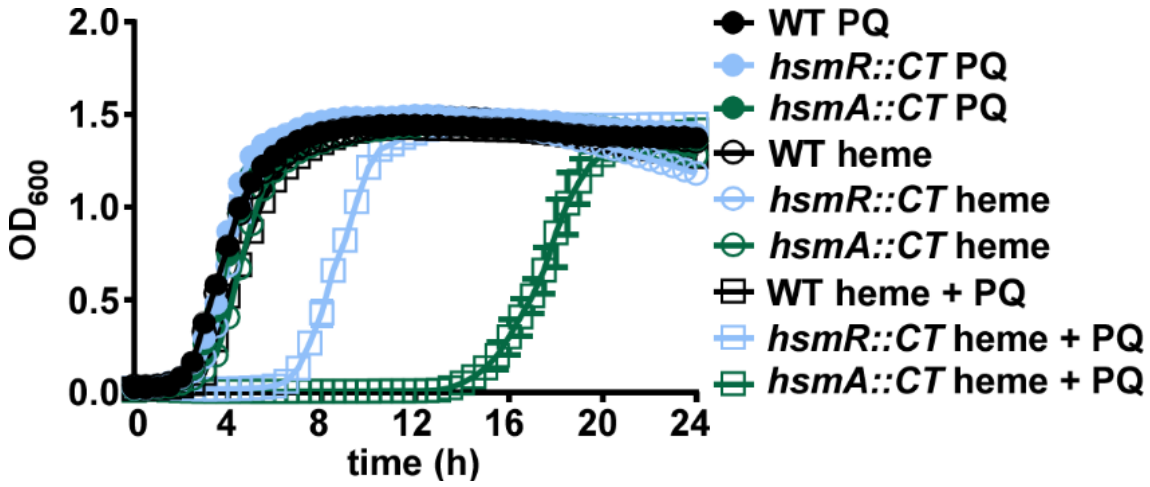

Figure 20 *hsmR::CT* and *hsmA::CT* are sensitive to paraquat in the presence of heme. Growth of *C. difficile* WT, *hsmR::CT*, and *hsmA::CT* in BHIS in the presence or absence of paraquat (2 mM) and heme (25 µM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

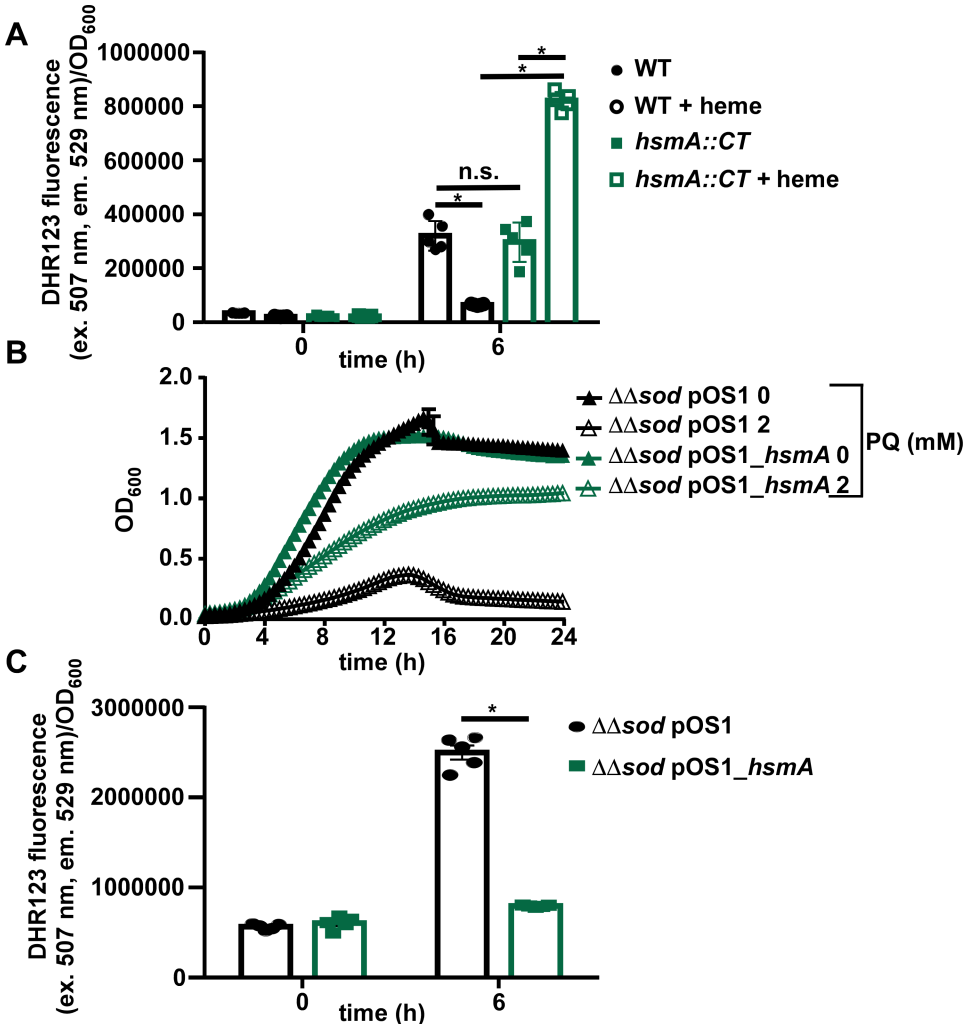

Figure 21 HsmA reduces oxidative stress.
(A) *C. difficile* WT and *hsmA::CT* strains were grown for 6 hours followed by treatment with or without heme (25 μM) for 30 min. Samples were exposed to atmospheric oxygen and oxidative stress generation was determined by measuring fluorescence of dihydrorhodamine 123 (ex. 507 nm, em. 529). (B) Growth of *S. aureus* ΔΔ*sod* pOS1 and pOS1_*hsmA* strains in the presence or absence of paraquat (2 mM) and dihydrorhodamine 123. (C) Oxidative stress generation was quantified by measuring fluorescence of dihydrorhodamine 123 (ex. 507 nm, em. 529). The data are a representative from three independent experiments each in biological quintuplicate with standard error of the mean. Statistical significance was determined using the multiple comparison two-way ANOVA test with the Sidak correction for multiple comparisons comparing the means of each group to one another. * denotes p-value < 0.05

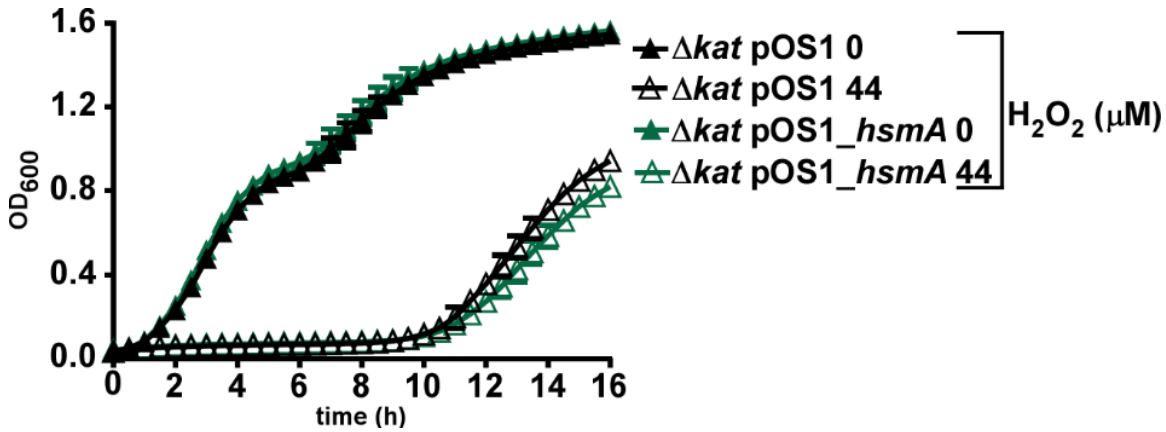

Figure 22 HsmA is not involved in protection against hydrogen peroxide.
Growth of *S. aureus* Δ*kat* pOS1 and pOS1_*hsmA* strains in the presence or absence of H_2O_2 (44 μM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

HsmR and HsmA promote resistance to vancomycin during CDI.

The ability of HsmA to reduce oxidative stress could protect against antibiotic toxicity, as it has previously been suggested that bactericidal antibiotics generate considerable oxidative stress resulting from hyper-induced metabolism [84]. To investigate the protective effects of the *hsmRA* operon against antibiotics, we determined the sensitivities of WT, *hsmR::CT,* and *hsmA::CT* to vancomycin as measured by growth over time. All three strains were equally sensitive to vancomycin treatment alone (Figure 23A-B). When heme was added in combination with vancomycin, the WT strain displayed a significant growth recovery, while the *hsmR::CT* strain exhibited a significant delay in growth and the *hsmA::CT* strain did not grow (Figure 23A-B). This growth recovery is specific to heme as PPIX did not rescue growth in the WT strain (Figure 24). Heme treatment additionally rescued growth of the WT strain but not *hsmR::CT* and *hsmA::CT* in the presence of metronidazole (Figure 25). These results demonstrate that HsmA reduces toxicity against redox damage produced directly or indirectly by clinically relevant glycopeptide and nitroimidazole antibiotics[84, 85].

C. difficile infection leads to a significant accumulation of heme in the lumen of the gut [65]. To investigate the involvement of the *hsmRA* operon in persistence during infection, mice were infected with a 1:1 ratio of WT and *hsmR::CT* or WT and *hsmA::CT* spores. Disease was monitored for 4 days, followed by a 5-day treatment of vancomycin (0.2 mg/mL), after which mice were monitored for relapse. During the acute phase of infection, all strains colonized the mice, as exhibited by the ~10^8 colony-forming units (CFU) of the WT strain recovered per gram of stool and the ~10^6 -10^7 CFU of the *hsmR::CT* and *hsmA::CT* strains recovered per gram of stool (Figure 23C-D). During vancomycin treatment, *hsmR::CT* and *hsmA::CT* colonization levels were below the limit of detection for the length of the antibiotic treatment in contrast to the temporal reduction

(~10^6 to ~10^2) in WT CFUs per gram of stool (Figure 23C-D). Upon removal of vancomycin, *hsmR::CT* and *hsmA::CT* exhibited a 2 – 3-log defect in recovery compared to the WT in the first day of relapse. On the final day of infection, the mutant strains displayed a 0.5 – 2-log reduction of CFUs per gram of cecal contents compared to WT (Figure 23C-D). These results demonstrate the *hsmRA* operon provides protection against vancomycin treatment in addition to being required for full relapse during CDI colonization. Taken together, these data establish HsmR as a sensor of host heme that induces the production of HsmA which binds heme and prevents toxicity associated with heme accumulation, while also providing protection against antibiotics and oxidative stress. Combined, these activities contribute to the ability of *C. difficile* to persist during infection despite the oxidative burst of phagocytes or antimicrobial treatment.

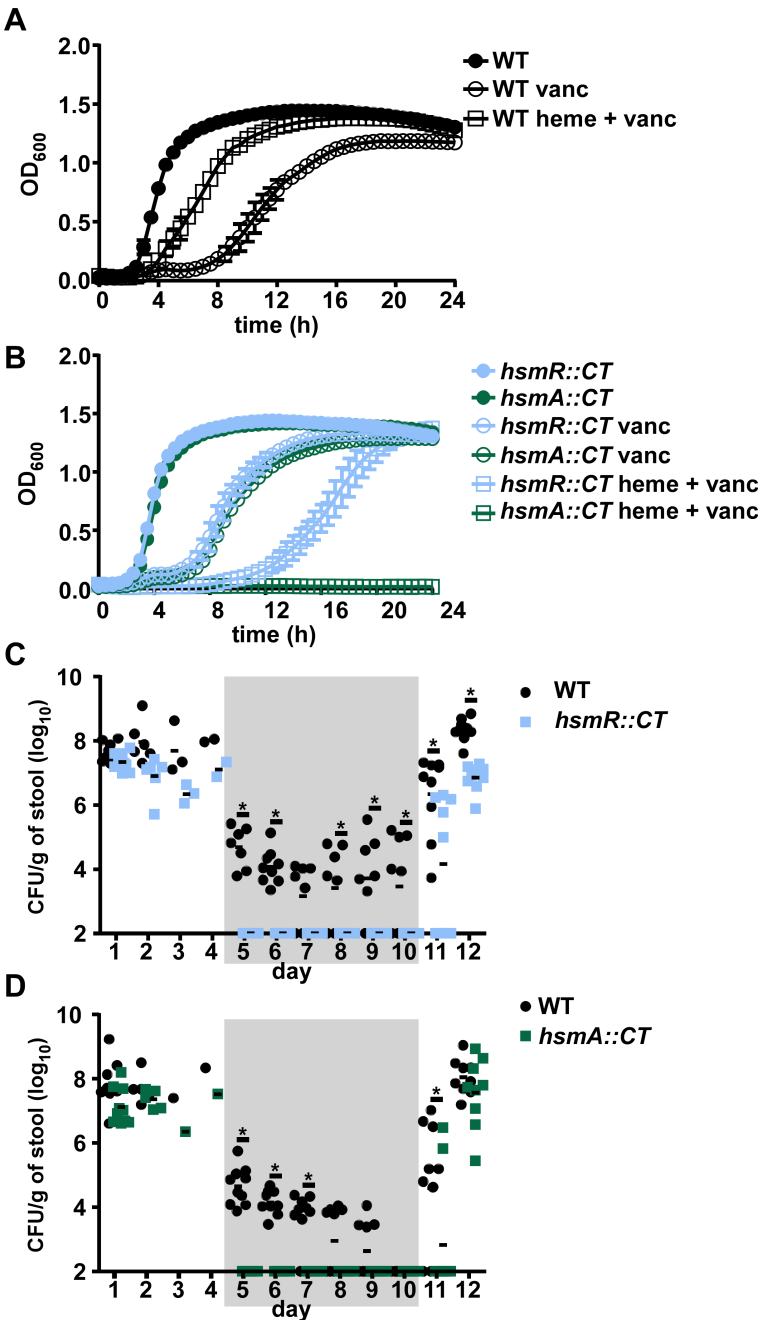

Figure 23 The *hsmRA* operon decreases sensitivity to vancomycin in the presence of heme during infection.

Growth of (A) WT, (B) *hsmR::CT* and (B) *hsmA::CT* in the presence or absence of vancomycin (100 µg/mL) and heme (10 µM). CFU analysis of mice co-infected with (C) WT and *hsmR::CT* or (D) WT and *hsmA::CT* strains with standard error on the mean (n = 10/group). Vancomycin treatment (0.4 mg/mL) was administered on days 5 – 10 (denoted by gray shading) and removed on day 11. The data are a representative from three independent experiments each in biological triplicate with standard error of the mean. Statistical significance was determined using the multiple comparison two-way ANOVA test with the Bonferroni correction for multiple comparisons comparing the means of each group to one another. * denotes p-value < 0.05

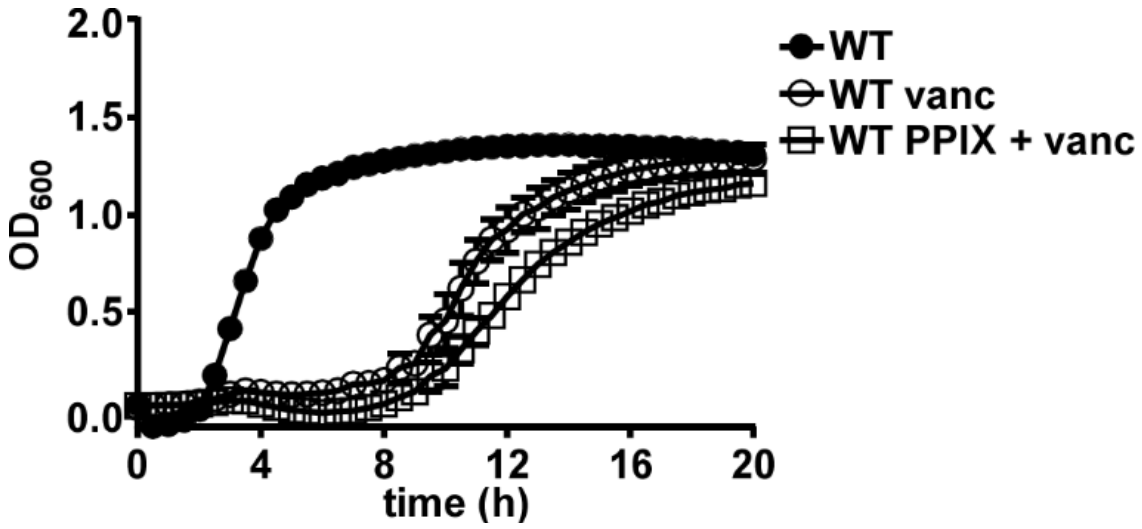

Figure 24 Protoporphyrin-IX does not rescue vancomycin toxicity.
Growth of WT in the presence or absence of vancomycin (100 g/mL) and PPIX (50 μM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

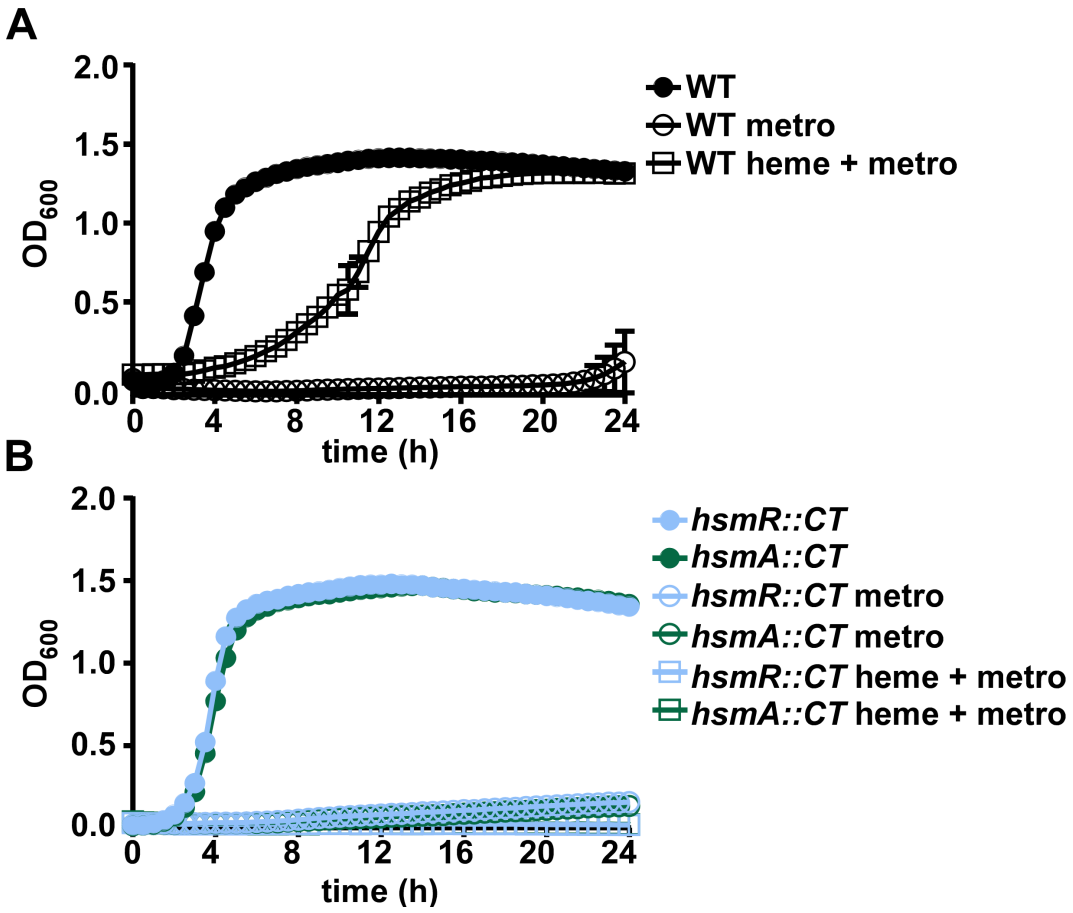

Figure 25 HsmRA decreases sensitivity to metronidazole in the presence of heme. Growth of (A) WT, (B) *hsmR::CT* and (B) *hsmA::CT* in the presence or absence of metronidazole (2 mg/mL) and heme (10 μM). The data are a representative from three independent experiments each in biological triplicate with standard error of the mean.

To predict the generalizability of HsmA-mediated protection against oxidative stress across the Bacterial kingdom, we investigated the conservation of HsmA across species. A phylogenetic tree was created using HsmA from *C. difficile* as a seed. This analysis revealed that HsmA is widespread in *Clostridia* and candidate HsmA orthologues can be found in multiple genera including *Bacteroides*, *Bacillus*, *Lactobacillus*, *Enterococcus*, and *Geobacter* (Figure 26, Appendix B Table 5), suggesting that HsmA may represent a conserved strategy for dealing with environmental oxidants across numerous organisms.

Figure 26 HsmA is conserved among multiple bacterial species.

Distribution of HsmA homologues represented by solid blue dots among 120 reference bacteria with an added 37 representative bacteria with complete genomes. The phylogenetic tree was generated using the phylogenetic tree tool of the Patric database as described in the methods section and visualized in iTol.

Discussion

C. difficile thrives in the colon during infection despite creating a hostile inflammatory environment through toxin-mediated damage of the gastrointestinal epithelium [7, 8, 86]. Robust inflammation leads to high levels of heme at the host-pathogen interface during CDI (Figure 1) [65]. Herein, I identified an acutely heme responsive transcription factor HsmR which activates expression of the membrane protein HsmA that incorporates the reactivity of heme to defend against redox stress while simultaneously detoxifying excess heme through sequestration.

In chapter II I identified the HatRT system in *C. difficile* that senses and detoxifies excess intracellular heme through efflux [65]. The results herein further refine my current model of heme homeostasis in *C. difficile*. HsmR senses low concentrations of heme and activates expression of the *hsmRA* operon which leads to the integration of heme into HsmA. Heme-bound HsmA within the membrane shields the bacterium against redox active molecules. After intracellular heme reaches a certain threshold that can no longer be utilized by HsmA, HatR binds heme, derepressing the *hatRT* operon, leading to subsequent efflux of heme through HatT, resulting in a relief from heme toxicity. Together these systems function to maintain an optimal concentration of intracellular heme for *C. difficile* to protect itself against the stressors of the host and achieve full pathogenicity.

The inflamed gut contains a plethora of environmental, microbiota, and host mediated stressors [9, 87, 88]. In addition to heme toxicity at the host pathogen interface, *C. difficile* as an obligate anaerobe encounters oxidative stress in various forms ranging from oxygenation of the colonic epithelium due to inflammation and ROS produced by host immune cells (Figure 1) [9, 89]. The HsmA-dependent decrease of oxidative stress may be the result of a serendipitous evolutionary event, as heme toxicity and oxidative stress coincide temporally. The mechanism by

which HsmA protects against oxidative stress remains unknown. HsmA may function to enzymatically convert a radical species into a harmless form or it may shuttle electrons through the bound heme to act as an electron sink, consistent with its homology to cytochrome b561. The ability of heme-bound HsmA to protect against different classes of antibiotics may be due to the elicitation of oxidative stress by bactericidal antibiotics [84, 90]. While *C. difficile* does not aerobically respire, the altered metabolism induced by antibiotics, as observed in other bacterial species, might generate oxidants that eventually lead to death [91]. This accumulation of oxidants is countered by heme-HsmA complexes, which reduce their concentrations below lethal limits. In total, coating of the membrane with heme-HsmA complexes provides a shield against redox damage produced in the inflamed gastrointestinal tract during infection.

A bioinformatic analysis of the *C. difficile* genome reveals an incomplete heme biosynthesis pathway, as there are no identified ferrochelatase or δ-amino-levulinic acid synthesis genes despite the presence of genes required for siroheme and cobalamin synthesis [37, 38]. These data indicate that *C. difficile* acquires heme exogenously during infection, presumably from the host due to toxin-mediated damage of the gastrointestinal epithelial layer. Additionally, as *C. difficile* cannot use heme as a sole iron source, the limited transcriptional response of *C. difficile* to heme exposure and lack of identified heme cofactor proteins suggests HsmA as the primary protein to utilize heme in this organism [39]. However, the mechanism by which heme enters *C. difficile* remains unknown. The phenomenon of exogenous heme utilization by bacteria that cannot synthesize endogenous heme has been observed [13]. In *Enterococcus faecalis*, exogenous heme induces the production of a heme containing catalase [14]. Numerous lactic acid bacteria, such as *Lactococcus lactis*, acquire exogenous heme to establish aerobic respiratory chains [15]. In this

chapter, I demonstrate *C. difficile* uses exogenous heme from the host as a cofactor for HsmA, which provides resistance against antimicrobial stressors.

Heme sequestration to reduce heme toxicity is a conserved strategy in multiple pathogenic organisms [20]. Most sequestration proteins have been described in Gram-negative pathogens and consist of intracellular heme binding proteins such as the HemS family in *Yersinia enterocolitica, Shigella dysenteriae, Pseudomonas aeruginosa*, and *E. coli* [31, 92-94]. These proteins often have additional functions dependent upon the organism such as storage, trafficking, or degradation but all contribute to heme detoxification [20]. HsmA is widespread in the Clostridial species, but it is not limited to this genus nor to anaerobes as orthologs are present in Bacteroidetes and Bacilli as well as the Proteobacteria (Figure 26). In several *Bacillus cereus* species, a physical clustering exists between genes encoding an ortholog of HsmA and the heme efflux pump HrtAB suggesting evolutionary pressure to genetically cluster heme detoxification systems in certain species (Figure 27) [25]. The heme detoxification proteins encoded by *hsmRA* represent a unique mechanism of sequestration that appears to be wide spread among bacteria that interact with vertebrate blood or environmental heme during their lifecycle.

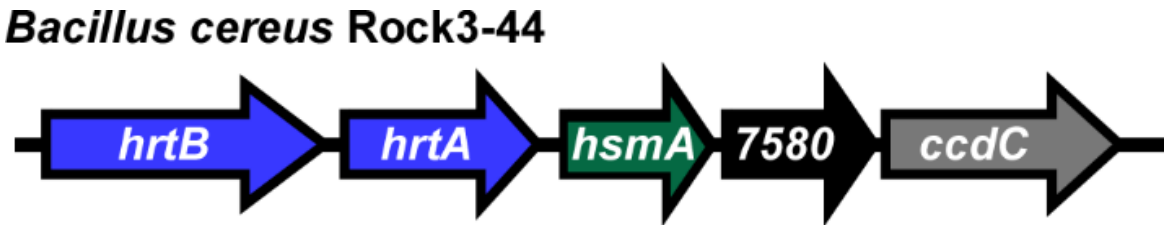

Figure 27 Physical clustering of genes encoding HsmA and the heme efflux proteins HtrAB. *Bacillus cereus* Rock3-44 *hsmA*, (NCBI protein accession number ZP_04216402.1*)* Abbreviations: *ccdC*, cytochrome c biogenesis protein CcdC; *hrtA*, Heme efflux system ATPase HrtA; *hrtB*, Heme efflux system permease HrtB.

Together, these results demonstrate that *C. difficile* HsmRA capitalizes on the toxin-induced inflammation of the gastrointestinal tract to seize heme from the host to protect against antibiotic therapy and immune cell mediated oxidative stress produced at the host-pathogen interface (Figure 1). The primary treatment of CDI in patients is a vancomycin regimen where 20% of patients have recurrent infection resulting in additional antibiotic treatment or, ultimately, a fecal microbiota transfer [6]. In a murine model of relapse CDI, the *hsmR::CT* and *hsmA::CT* mutant strains were significantly more sensitive to vancomycin treatment. These data provide support for the development of a drug targeting HsmR or HsmA as a therapy for CDI that can be used in combination with these antibiotics. Further studies will elucidate the biochemical mechanisms of HsmA induced protection against oxidative stress in *C. difficile* and potentially other organisms.

CHAPTER IV

SUMMARY AND FUTURE DIRECTIONS

Heme sensing, utilization, and detoxification in *C. difficile*

C. difficile is a formidable gastrointestinal pathogen that has evolved an arsenal of mechanisms to thrive in the host during infection [4, 9]. Upon colonization of the colon, *C. difficile* causes severe damage to the gastrointestinal epithelium through the production of the toxins TcdA and TcdB [4]. The resulting damage causes perforations in the gut leading to fluid secretion, translocation of erythrocytes into the lumen, and neutrophil recruitment to the site of infection [9]. Despite the antimicrobial molecules and stressors produced at the host-pathogen interface, *C. difficile* disease can progress ultimately leading to death [3]. Understanding the mechanisms *C. difficile* utilizes to survive the harsh environment of the host may prove to be critical to the development of future therapies towards this urgent public health threat.

A primary focus of my dissertation was to demonstrate *C. difficile* encounters heme stress during infection and identify the molecular mechanisms this pathogen utilizes to maintain heme homeostasis. In chapter II, I employed MALDI-IMS to visualize a high concentration of hemoglobin and therefore heme at sites of inflammation and damage in the murine cecum as a result of *C. difficile* infection. I further identified heme is toxic to *C. difficile* and *C. difficile* adapts to heme toxicity if exposed to a non-lethal heme concentration prior to exposure to higher concentrations. To reveal the genes encoding for proteins involved in this adaptation, I performed an RNA-seq comparing *C. difficile* treated with a sub-toxic concentration of heme to an untreated control. These data led to the identification and naming of the *hatRT* operon. This operon encodes for a TetR-family transcriptional regulator (HatR) and a major facilitator superfamily transporter

(HatT). The gene products encoded by this operon sense heme by HatR binding and detoxify excess intracellular heme by HatT mediated efflux. To my knowledge this is the first identified heme efflux system described in an obligate anaerobe. A strain with *hatT* inactivated was further identified to have a toxin-independent reduction in pathogenicity in a mouse model of CDI establishing the importance of heme detoxification during infection. Despite this reduction in pathogenicity with the absence of *hatT*, *C. difficile* strains fully colonize the mice which led to the hypothesis that *C. difficile* contains additional mechanisms of heme detoxification.

In chapter III, I returned to the RNA-seq performed in chapter II to identify other candidate genes involved in heme homeostasis. I identified and named the two most transcriptionally increased genes the *hsmRA* operon. This operon encodes for a MarR-family transcriptional regulator (HsmR) and a membrane protein with homology to a cytochrome b561 (HsmA). I discovered HsmR senses heme through binding which activates the expression of the *hsmRA* operon where HsmA reduces heme toxicity through sequestration. Furthermore, I revealed the transcriptional response of *C. difficile* to a brief exposure of heme limited to the *hsmRA* and *hatRT* operons. Given the homology of HsmA to cytochromes, I hypothesized that HsmA uses heme as a cofactor for other important functions. As *C. difficile* does not respire and cannot use heme as a sole iron source, I investigated if heme-bound HsmA protects against redox active molecules. I discovered when the WT strain was treated with heme there was a protective effect against ROS whereas strains lacking either *hsmR* or *hsmA* displayed an increased sensitive to redox stress in the presence of heme. Expressing HsmA in a strain of *S. aureus* lacking both SOD enzymes revealed a reduction in detected oxidative stress when treated with the superoxide generating molecule paraquat. With the theory that many antibiotics kill bacteria through the generation of ROS, I investigated if HsmA provided protection against the clinically relevant antibiotics vancomycin

and metronidazole. I discovered that in the presence of heme *C. difficile* is more resistant to both antibiotics and this protection is mediated by HsmA. The protection provided by HsmA extended to a murine model of relapse infection as strains with *hsmR* or *hsmA* inactivated have increased sensitivity to vancomycin treatment and reduced colonization persistence. My results define a mechanism exploited by *C. difficile* to repurpose toxic heme liberated within the inflamed gastrointestinal tract as a shield against antimicrobial compounds encountered at the host-pathogen interface.

With the results herein I developed our current model of heme homeostasis in *C. difficile*. HsmR senses low concentrations of heme and activates expression of the *hsmRA* operon which leads to the integration of heme into HsmA. Heme bound HsmA within the membrane shields the bacterium against redox active molecules. After intracellular heme reaches a certain threshold that can no longer be utilized by HsmA, HatR binds heme depressing the *hatRT* operon leading to subsequent efflux of heme through HatT resulting in a relief from heme toxicity. Together these systems function to maintain an optimal concentration of intracellular heme for *C. difficile* to protect itself against the stressors of the host and achieve full pathogenicity (Figure 28). My results set the stage for the development of therapeutic interventions to target these bacterial-specific systems.

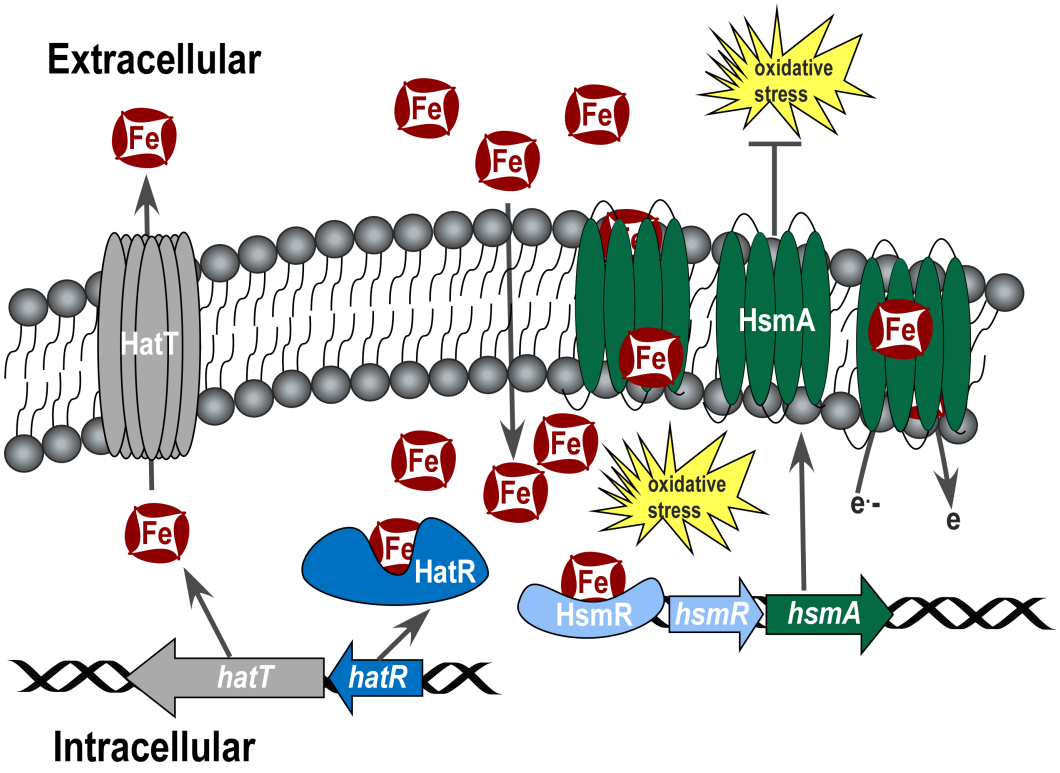

Figure 28 Heme sensing, utilization, and detoxification in *C. difficile*. Host heme is sensed by HsmR and incorporated into HsmA, providing protection against oxidative stress produced by host immune cells and environment. Upon reaching toxic intracellular concentrations, HatR binds heme depressing the *hatRT* operon and leading to subsequent efflux of heme by HatT.

Defining anaerobic heme toxicity in *C. difficile*

Many aspects of heme entry and toxicity remain unknown in *C. difficile*. My work presented in this dissertation demonstrated the ability of heme to enter the cytoplasm of *C. difficile* but did not identify a dedicated mechanism of import. The sole genes upregulated to a brief exposure of heme were the *hsmRA* and *hatRT* operons suggesting if a dedicated mechanism of transport exists, it may be through a non-specific transporter or does not display increased transcription in the presence of heme. In order to test whether the accumulation of heme within the cytoplasm is due to passive diffusion of heme through the membrane or through a protein mediated mechanism, a transposon sequencing experiment utilizing a heme sensitive screening should be employed. If the mechanism is passive diffusion through the membrane, these data may suggest a primary mediator of toxicity would result from membrane damage and disruption. Furthermore, an understanding of heme toxicity in an anaerobic environment is far from complete and could be gained from additional investigation into the transcriptional response of *C. difficile* to heme. In chapters II and III, I provided data from two RNA-seq experiments differing in the amount of time *C. difficile* was exposed to heme. A brief exposure of heme solely increased the transcription of 4 genes compared to the 245 transcriptionally increased genes when constantly exposed to heme throughout growth. This lends to the identification of secondary pathways that are activated in response to persistent heme exposure. Additional RNA-seq experiments or targeted transcriptional analysis of intermediate time points of exposure could reveal secondary responsive genes as well as identify mechanisms of anaerobic heme toxicity.

Structural and mechanistic investigation of HatRT and HsmRA

In this work I have assigned functions of heme sensing, utilization, and detoxification to the HatRT and HsmRA protein systems however, the enzymatic and biochemical mechanisms of these functions are not fully defined. Further investigation is required to understand how HatR and HsmR coordinate heme and the resulting modification in structure after binding that allows for alteration in DNA-binding affinity. Additionally, investigations into the fates of these regulators after binding heme may reveal other roles outside of regulation such as heme trafficking to either HatT or HsmA.

As with most heme efflux transporters, it is not known if HatT transports heme or some derivative thereof. The heme could be contained in a protein, trapped within the membrane, or free in the cytoplasm. Furthermore, the mechanism of heme efflux of HatT is unknown. HatT is a member of the major facilitator superfamily of transporters, which primarily utilize a rocker-switch mechanism to facilitate the transfer of a molecule from one side of a membrane to another which is dependent on specific residues in the binding cavity [95]. Investigations into the involved residues would elucidate the specificity and reveal the mechanism of efflux. The generation of a crystal or Cryo-EM structure in the presence and absence of heme would provide an understand of how this transporter functions and provide the basis for additional investigations in other potential interacting proteins.

The ability of HsmA to reduce oxidative stress provides a potential paradigm of a heme utilizing protein that protect against oxidative stress in bacteria. The mechanism of protection has not been defined. I demonstrated in chapter III that HsmA binds heme however, how the bound heme interacts with oxidative species was not determined. Specific analytical assays involving electrochemical probes and artificial membranes could be utilized with the purified recombinant

HsmA to reveal the flow of electrons and if enzymatic products are produced. Furthermore, identifying the specific residues involved with heme binding and potential enzymatic activity would mechanistically reveal how HsmA protects against redox active molecules. These studies could be expanded to other clostridial and bacterial species that contain HsmA orthologs to demonstrate this protein as a conserved mechanism to reduce oxidative stress. Taken together, gaining a greater understanding of how HatRT and HsmRA function at a structural and molecular level could provide ideal targets for therapeutic design.

Further elucidation of the contribution of HatRT and HsmRA to *C. difficile* pathogenicity

The investigation of heme sensing, utilization, and detoxification in *C. difficile* pathogenicity presented in this work demonstrated the importance of these systems during infection. However, I was unable to fully define the contribution of both HatRT and HsmRA systems to *C. difficile* pathogenicity as I was unsuccessful at creating a strain with HatT and HsmA inactivated due to the limited ability of current genetic tools for *C. difficile*. Advanced genetic tools are currently under development that may lead to the generation of a *hatT* and *hsmR* double mutant strain. I hypothesize that this strain would be severely defective in persistence and display significantly reduced pathogenicity in the inflamed gastrointestinal tract in the acute murine model of CDI. Moreover, the strain lacking *hsmA* displays increased sensitivity to vancomycin therefore I hypothesize that a double mutant would be unable to recover after antibiotic treatment in the relapse murine model of CDI. An investigation into different models of infection, such as the golden Syrian hamster model, could reveal additional contributions of HatT and HsmA to infection.

A strain lacking HatT was able to fully colonize the inflamed gastrointestinal tract at a cost of reduced pathogenicity without a reduction in toxin titer. This observation suggests that either *C. difficile* directly alters its pathogenicity to compensate for the inability to fully detoxify heme or that the inability to efflux heme changes the environment or the microbiota during infection. To investigate the hypothesis that *C. difficile* directly modifies pathogenicity, a NanoString assay with a specific probes to investigate virulence, motility, and stress response proteins on infected cecal contents comparing *hatT::CT* to the WT strain would reveal transcriptional modifications during infection. Furthermore, the physical location of *C. difficile* specifically regarding the proximity to the gastrointestinal epithelial cells has not previously been defined. Investigations of the physical space that *hatT::CT* occupies utilizing fluorescent microscopy of infected mouse ceca with probes specific for *C. difficile* could demonstrate *C. difficile* reduces the potential exposure to heme by distancing itself from the sites of damage. This may also account for the reduction in observed pathology despite the consistent level of toxin being produced. Alternatively, heme is a sought-after resource of many different bacterial species and if HatT transports intact heme, I hypothesize the inability to efflux heme would shape the structure of the microbiota. A 16s sequencing experiment of mice infected with the *hatT::CT* strain compared to mice infected with the WT strain may reveal significantly different distributions of the microbiota. As the composition of the microbiota often directly correlates to disease state, the reduction of this valuable resource may prevent harmful members of the microbiota from blooming and increasing inflammation. Another hypothesis is the host is reacting to the free heme that is transported by HatT and induces additional inflammation. Luminex or other analysis of cytokines and chemokines produced in the *hatT::CT* mouse cecum compared to mice infected with the WT strain would identify a differential host response that is a cause of altered pathology.

In chapter III I demonstrated the ability of HsmA to reduce oxidative stress in the presence of heme. Previous studies have shown that the strain used in this work, R20291, is more sensitive to oxygen concentrations than other *C. difficile* strains such as 630 [96]. However, none of these studies investigated this sensitivity after the bacteria had been exposed to heme. A hypothesis is that the observed strain to strain differences after exposure to oxygen is dependent on the presence of heme and HsmA. To test this, multiple strains of *C. difficile* would be treated with heme prior to exposure of atmospheric oxygen and their survival and cellular oxidative stress measured. This assay could be expanded to use strains obtained from patients at Vanderbilt to understand if this utilization of heme to protect against oxidative stress is a universal phenomenon in *C. difficile*. Additionally, if there are variations between different patient isolates these data could be linked to the severity of infection observed in the patients. During *C. difficile* infection the oxygen levels in the gastrointestinal tract increase with the severity of inflammation. In chapter III, I tested the ability of HsmA to reduce redox active molecules in an anaerobic environment as well as in reduced media exposed to atmospheric oxygen. Considering *C. difficile* would experience high concentrations of heme after toxin mediated inflammation has begun to oxygenate the gut, I hypothesize that HsmA has evolved to optimally function at a certain concentration of oxygen. To test this, I would measure survival and cellular oxidative stress of *C. difficile* treated with heme at varying concentrations of oxygen in a hypoxia chamber. Furthermore, I would perform these assays in the presence of antibiotics and other redox generating molecules to elucidate the maximal protection provided by HsmA. In chapter III, I discuss how HsmA may additionally provide protection against immune mediators during infection. This ability was not directly investigated in this work. However, neutrophil killing assays have not been successfully designed for R20291 due to the high sensitivity of this strain to oxygen. If the observed protection of oxidative stress by

HsmA extends to more aerotolerant strains, I hypothesize treating *C. difficile* with heme would increase the resistance to killing by neutrophils. These data would provide a facet of the ability of *C. difficile* to maintain high bacterial densities in the inflamed colon despite the attempt of pathogen clearance by innate immune cells. Taken together, further defining the contribution of HatRT and HsmRA to *C. difficile* pathogenesis may reveal critical aspects of infection and an overall understanding of how *C. difficile* survives in the inflamed gastrointestinal tract during infection.

References

1. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of *Clostridium difficile* infection in the United States. N Engl J Med. 2015;372(9):825-34. doi: 10.1056/NEJMoa1408913. PubMed PMID: 25714160.

2. Abt MC, McKenney PT, Pamer EG. *Clostridium difficile* colitis: pathogenesis and host defence. Nat Rev Microbiol. 2016. doi: 10.1038/nrmicro.2016.108. PubMed PMID: 27573580.

3. Rupnik M, Wilcox MH, Gerding DN. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol. 2009;7(7):526-36. doi: 10.1038/nrmicro2164. PubMed PMID: 19528959.

4. Shen A. *Clostridium difficile* toxins: mediators of inflammation. J Innate Immun. 2012;4(2):149-58. Epub 2012/01/10. doi: 10.1159/000332946. PubMed PMID: 22237401; PubMed Central PMCID: PMCPMC3388264.

5. Guery B, Galperine T, Barbut F. *Clostridioides difficile*: diagnosis and treatments. BMJ. 2019;366:l4609. Epub 2019/08/20. doi: 10.1136/bmj.l4609. PubMed PMID: 31431428.

6. Cammarota G, Ianiro G, Gasbarrini A. Fecal microbiota transplantation for the treatment of *Clostridium difficile* infection: a systematic review. J Clin Gastroenterol. 2014;48(8):693-702. doi: 10.1097/MCG.0000000000000046. PubMed PMID: 24440934.

7. Carter GP, Chakravorty A, Pham Nguyen TA, Mileto S, Schreiber F, Li L, et al. Defining the Roles of TcdA and TcdB in Localized Gastrointestinal Disease, Systemic Organ Damage, and the Host Response during *Clostridium difficile* Infections. MBio. 2015;6(3):e00551. doi: 10.1128/mBio.00551-15. PubMed PMID: 26037121; PubMed Central PMCID: PMCPMC4453007.

8. Chumbler NM, Farrow MA, Lapierre LA, Franklin JL, Haslam DB, Haslam D, et al. *Clostridium difficile* Toxin B causes epithelial cell necrosis through an autoprocessing-independent mechanism. PLoS Pathog. 2012;8(12):e1003072. doi: 10.1371/journal.ppat.1003072. PubMed PMID: 23236283; PubMed Central PMCID: PMCPMC3516567.

9. Péchiné S, Collignon A. Immune responses induced by *Clostridium difficile*. Anaerobe. 2016;41:68-78. Epub 2016/04/21. doi: 10.1016/j.anaerobe.2016.04.014. PubMed PMID: 27108093.

10. Ahlquist DA, McGill DB, Schwartz S, Taylor WF, Owen RA. Fecal blood levels in health and disease. A study using HemoQuant. N Engl J Med. 1985;312(22):1422-8. doi: 10.1056/NEJM198505303122204. PubMed PMID: 3873009.

11. Aich A, Freundlich M, Vekilov PG. The free heme concentration in healthy human erythrocytes. Blood Cells Mol Dis. 2015;55(4):402-9. Epub 2015/09/21. doi: 10.1016/j.bcmd.2015.09.003. PubMed PMID: 26460266; PubMed Central PMCID: PMCPMC4860002.

12. Ng LG, Ostuni R, Hidalgo A. Heterogeneity of neutrophils. Nat Rev Immunol. 2019;19(4):255-65. doi: 10.1038/s41577-019-0141-8. PubMed PMID: 30816340.

13. Yamamoto Y, Poyart C, Trieu-Cuot P, Lamberet G, Gruss A, Gaudu P. Respiration metabolism of Group B Streptococcus is activated by environmental haem and quinone and contributes to virulence. Mol Microbiol. 2005;56(2):525-34. doi: 10.1111/j.1365-2958.2005.04555.x. PubMed PMID: 15813741.

14. Frankenberg L, Brugna M, Hederstedt L. *Enterococcus faecalis* heme-dependent catalase. J Bacteriol. 2002;184(22):6351-6. doi: 10.1128/jb.184.22.6351-6356.2002. PubMed PMID: 12399505; PubMed Central PMCID: PMCPMC151946.

15. Lechardeur D, Cesselin B, Fernandez A, Lamberet G, Garrigues C, Pedersen M, et al. Using heme as an energy boost for lactic acid bacteria. Curr Opin Biotechnol. 2011;22(2):143-9. Epub 2011/01/04. doi: 10.1016/j.copbio.2010.12.001. PubMed PMID: 21211959.

16. Mazmanian SK, Skaar EP, Gaspar AH, Humayun M, Gornicki P, Jelenska J, et al. Passage of heme-iron across the envelope of *Staphylococcus aureus*. Science. 2003;299(5608):906-9. doi: 10.1126/science.1081147. PubMed PMID: 12574635.

17. Maresso AW, Garufi G, Schneewind O. *Bacillus anthracis* secretes proteins that mediate heme acquisition from hemoglobin. PLoS Pathog. 2008;4(8):e1000132. doi: 10.1371/journal.ppat.1000132. PubMed PMID: 18725935; PubMed Central PMCID: PMCPMC2515342.

18. Létoffé S, Nato F, Goldberg ME, Wandersman C. Interactions of HasA, a bacterial haemophore, with haemoglobin and with its outer membrane receptor HasR. Mol Microbiol. 1999;33(3):546-55. PubMed PMID: 10417645.

19. Lewis LA, Gray E, Wang YP, Roe BA, Dyer DW. Molecular characterization of *hpuAB*, the haemoglobin-haptoglobin-utilization operon of *Neisseria meningitidis*. Mol Microbiol. 1997;23(4):737-49. PubMed PMID: 9157245.

20. Choby JE, Skaar EP. Heme Synthesis and Acquisition in Bacterial Pathogens. J Mol Biol. 2016. doi: 10.1016/j.jmb.2016.03.018. PubMed PMID: 27019298.

21. Imlay JA, Chin SM, Linn S. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. Science. 1988;240(4852):640-2. PubMed PMID: 2834821.

22. Wakeman CA, Hammer ND, Stauff DL, Attia AS, Anzaldi LL, Dikalov SI, et al. Menaquinone biosynthesis potentiates haem toxicity in *Staphylococcus aureus*. Mol Microbiol. 2012;86(6):1376-92. Epub 2012/10/24. doi: 10.1111/mmi.12063. PubMed PMID: 23043465; PubMed Central PMCID: PMCPMC3524387.

23. Nitzan Y, Wexler HM, Finegold SM. Inactivation of anaerobic bacteria by various photosensitized porphyrins or by hemin. Curr Microbiol. 1994;29(3):125-31. PubMed PMID: 7765091.

24. Nir U, Ladan H, Malik Z, Nitzan Y. In vivo effects of porphyrins on bacterial DNA. J Photochem Photobiol B. 1991;11(3-4):295-306. PubMed PMID: 1816365.

25. Torres VJ, Stauff DL, Pishchany G, Bezbradica JS, Gordy LE, Iturregui J, et al. A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. Cell Host Microbe. 2007;1(2):109-19. doi: 10.1016/j.chom.2007.03.001. PubMed PMID: 18005689; PubMed Central PMCID: PMCPMC2083280.

26. Stauff DL, Skaar EP. *Bacillus anthracis* HssRS signalling to HrtAB regulates haem resistance during infection. Mol Microbiol. 2009;72(3):763-78. doi: 10.1111/j.1365-2958.2009.06684.x. PubMed PMID: 19400785; PubMed Central PMCID: PMCPMC2891670.

27. Lechardeur D, Cesselin B, Liebl U, Vos MH, Fernandez A, Brun C, et al. Discovery of intracellular heme-binding protein HrtR, which controls heme efflux by the conserved HrtB-HrtA transporter in *Lactococcus lactis*. J Biol Chem. 2012;287(7):4752-8. doi: 10.1074/jbc.M111.297531. PubMed PMID: 22084241; PubMed Central PMCID: PMCPMC3281666.

28. Fernandez A, Lechardeur D, Derré-Bobillot A, Couvé E, Gaudu P, Gruss A. Two coregulated efflux transporters modulate intracellular heme and protoporphyrin IX availability in

Streptococcus agalactiae. PLoS Pathog. 2010;6(4):e1000860. doi: 10.1371/journal.ppat.1000860. PubMed PMID: 20421944; PubMed Central PMCID: PMCPMC2858704.

29. Maharshak N, Ryu HS, Fan TJ, Onyiah JC, Schulz S, Otterbein SL, et al. *Escherichia coli* heme oxygenase modulates host innate immune responses. Microbiol Immunol. 2015;59(8):452-65. doi: 10.1111/1348-0421.12282. PubMed PMID: 26146866; PubMed Central PMCID: PMCPMC4582649.

30. Lansky IB, Lukat-Rodgers GS, Block D, Rodgers KR, Ratliff M, Wilks A. The cytoplasmic heme-binding protein (PhuS) from the heme uptake system of *Pseudomonas aeruginosa* is an intracellular heme-trafficking protein to the delta-regioselective heme oxygenase. J Biol Chem. 2006;281(19):13652-62. doi: 10.1074/jbc.M600824200. PubMed PMID: 16533806.

31. Wyckoff EE, Lopreato GF, Tipton KA, Payne SM. *Shigella dysenteriae* ShuS promotes utilization of heme as an iron source and protects against heme toxicity. J Bacteriol. 2005;187(16):5658-64. doi: 10.1128/JB.187.16.5658-5664.2005. PubMed PMID: 16077111; PubMed Central PMCID: PMCPMC1196095.

32. Stauff DL, Torres VJ, Skaar EP. Signaling and DNA-binding activities of the *Staphylococcus aureus* HssR-HssS two-component system required for heme sensing. J Biol Chem. 2007;282(36):26111-21. doi: 10.1074/jbc.M703797200. PubMed PMID: 17635909.

33. Hagman KE, Pan W, Spratt BG, Balthazar JT, Judd RC, Shafer WM. Resistance of *Neisseria gonorrhoeae* to antimicrobial hydrophobic agents is modulated by the mtrRCDE efflux system. Microbiology. 1995;141 (Pt 3):611-22. doi: 10.1099/13500872-141-3-611. PubMed PMID: 7711899.

34. Thompson JM, Jones HA, Perry RD. Molecular characterization of the hemin uptake locus (*hmu*) from *Yersinia pestis* and analysis of *hmu* mutants for hemin and hemoprotein utilization. Infect Immun. 1999;67(8):3879-92. PubMed PMID: 10417152; PubMed Central PMCID: PMCPMC96668.

35. Skaar EP, Gaspar AH, Schneewind O. IsdG and IsdI, heme-degrading enzymes in the cytoplasm of *Staphylococcus aureus*. J Biol Chem. 2004;279(1):436-43. doi: 10.1074/jbc.M307952200. PubMed PMID: 14570922.

36. LaMattina JW, Nix DB, Lanzilotta WN. Radical new paradigm for heme degradation in *Escherichia coli* O157:H7. Proc Natl Acad Sci U S A. 2016;113(43):12138-43. Epub 2016/10/10. doi: 10.1073/pnas.1603209113. PubMed PMID: 27791000; PubMed Central PMCID: PMCPMC5087033.

37. Moore SJ, Warren MJ. The anaerobic biosynthesis of vitamin B12. Biochem Soc Trans. 2012;40(3):581-6. doi: 10.1042/BST20120066. PubMed PMID: 22616870.

38. Dailey HA, Dailey TA, Gerdes S, Jahn D, Jahn M, O'Brian MR, et al. Prokaryotic Heme Biosynthesis: Multiple Pathways to a Common Essential Product. Microbiol Mol Biol Rev. 2017;81(1). Epub 2017/01/25. doi: 10.1128/MMBR.00048-16. PubMed PMID: 28123057; PubMed Central PMCID: PMCPMC5312243.

39. Cernat RC, Scott KP. Evaluation of novel assays to assess the influence of different iron sources on the growth of *Clostridium difficile*. Anaerobe. 2012;18(3):298-304. doi: 10.1016/j.anaerobe.2012.04.007. PubMed PMID: 22554901.

40. Jenior ML, Leslie JL, Young VB, Schloss PD. *Clostridium difficile* Colonizes Alternative Nutrient Niches during Infection across Distinct Murine Gut Microbiomes. mSystems.

2017;2(4). Epub 2017/07/25. doi: 10.1128/mSystems.00063-17. PubMed PMID: 28761936; PubMed Central PMCID: PMCPMC5527303.

41. Cartman ST, Minton NP. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. Appl Environ Microbiol. 2010;76(4):1103-9. Epub 2009/12/18. doi: 10.1128/AEM.02525-09. PubMed PMID: 20023081; PubMed Central PMCID: PMCPMC2820977.

42. Francis MB, Allen CA, Shrestha R, Sorg JA. Bile acid recognition by the *Clostridium difficile* germinant receptor, CspC, is important for establishing infection. PLoS Pathog. 2013;9(5):e1003356. doi: 10.1371/journal.ppat.1003356. PubMed PMID: 23675301; PubMed Central PMCID: PMCPMC3649964.

43. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102. doi: 10.1186/gb-2009-10-9-r102. PubMed PMID: 19781061; PubMed Central PMCID: PMCPMC2768977.

44. Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 1983;166(4):557-80. PubMed PMID: 6345791.

45. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. Science. 1997;277(5331):1453-62. PubMed PMID: 9278503.

46. Jeong H, Barbe V, Lee CH, Vallenet D, Yu DS, Choi SH, et al. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). J Mol Biol. 2009;394(4):644-52. Epub 2009/09/26. doi: 10.1016/j.jmb.2009.09.052. PubMed PMID: 19786035.

47. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.

48. Nairn BL, Lonergan ZR, Wang J, Braymer JJ, Zhang Y, Calcutt MW, et al. The Response of *Acinetobacter baumannii* to Zinc Starvation. Cell Host Microbe. 2016;19(6):826-36. doi: 10.1016/j.chom.2016.05.007. PubMed PMID: 27281572; PubMed Central PMCID: PMCPMC4901392.

49. Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. 2010;11:395. Epub 2010/07/23. doi: 10.1186/1471-2105-11-395. PubMed PMID: 20650010; PubMed Central PMCID: PMCPMC2918584.

50. Winston JA, Thanissery R, Montgomery SA, Theriot CM. Cefoperazone-treated Mouse Model of Clinically-relevant *Clostridium difficile* Strain R20291. J Vis Exp. 2016;(118). Epub 2016/12/10. doi: 10.3791/54850. PubMed PMID: 28060346; PubMed Central PMCID: PMCPMC5226375.

51. Zackular JP, Moore JL, Jordan AT, Juttukonda LJ, Noto MJ, Nicholson MR, et al. Dietary zinc alters the microbiota and decreases resistance to *Clostridium difficile* infection. Nat Med. 2016. doi: 10.1038/nm.4174. PubMed PMID: 27668938.

52. Theriot CM, Koumpouras CC, Carlson PE, Bergin II, Aronoff DM, Young VB. Cefoperazone-treated mice as an experimental platform to assess differential virulence of *Clostridium difficile* strains. Gut Microbes. 2011;2(6):326-34. Epub 2011/11/01. doi: 10.4161/gmic.19142. PubMed PMID: 22198617; PubMed Central PMCID: PMCPMC3337121.

53. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Anal Chem. 1997;69(23):4751-60. PubMed PMID: 9406525.

54. Moore JL, Caprioli RM, Skaar EP. Advanced mass spectrometry technologies for the study of microbial pathogenesis. Curr Opin Microbiol. 2014;19:45-51. doi: 10.1016/j.mib.2014.05.023. PubMed PMID: 24997399; PubMed Central PMCID: PMCPMC4125470.

55. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus Clostridium. J Microbiol Methods. 2007;70(3):452-64. doi: 10.1016/j.mimet.2007.05.021. PubMed PMID: 17658189.

56. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang X, et al. The TetR family of transcriptional repressors. Microbiol Mol Biol Rev. 2005;69(2):326-56. doi: 10.1128/MMBR.69.2.326-356.2005. PubMed PMID: 15944459; PubMed Central PMCID: PMCPMC1197418.

57. Stryer L. A conformation-dependent Cotton effect in the Soret band of hemin:poly-L-lysine. Biochim Biophys Acta. 1961;54:395-7. PubMed PMID: 13917913.

58. Awad MM, Johanesen PA, Carter GP, Rose E, Lyras D. *Clostridium difficile* virulence factors: Insights into an anaerobic spore-forming pathogen. Gut Microbes. 2014;5(5):579-93. doi: 10.4161/19490976.2014.969632. PubMed PMID: 25483328; PubMed Central PMCID: PMCPMC4615314.

59. Stauff DL, Bagaley D, Torres VJ, Joyce R, Anderson KL, Kuechenmeister L, et al. *Staphylococcus aureus* HrtA is an ATPase required for protection against heme toxicity and

prevention of a transcriptional heme stress response. J Bacteriol. 2008;190(10):3588-96. doi: 10.1128/JB.01921-07. PubMed PMID: 18326576; PubMed Central PMCID: PMCPMC2395006.

60. Sawai H, Yamanaka M, Sugimoto H, Shiro Y, Aono S. Structural basis for the transcriptional regulation of heme homeostasis in *Lactococcus lactis*. J Biol Chem. 2012;287(36):30755-68. doi: 10.1074/jbc.M112.370916. PubMed PMID: 22798069; PubMed Central PMCID: PMCPMC3436319.

61. Wakeman CA, Stauff DL, Zhang Y, Skaar EP. Differential activation of *Staphylococcus aureus* heme detoxification machinery by heme analogues. J Bacteriol. 2014;196(7):1335-42. doi: 10.1128/JB.01067-13. PubMed PMID: 24443529; PubMed Central PMCID: PMCPMC3993332.

62. Nobles CL, Green SI, Maresso AW. A product of heme catabolism modulates bacterial function and survival. PLoS Pathog. 2013;9(7):e1003507. Epub 2013/07/25. doi: 10.1371/journal.ppat.1003507. PubMed PMID: 23935485; PubMed Central PMCID: PMCPMC3723568.

63. Wilks A, Schmitt MP. Expression and characterization of a heme oxygenase (Hmu O) from *Corynebacterium diphtheriae*. Iron acquisition requires oxidative cleavage of the heme macrocycle. J Biol Chem. 1998;273(2):837-41. PubMed PMID: 9422739.

64. Tenhunen R, Marver HS, Schmid R. The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase. Proc Natl Acad Sci U S A. 1968;61(2):748-55. PubMed PMID: 4386763; PubMed Central PMCID: PMCPMC225223.

65. Knippel RJ, Zackular JP, Moore JL, Celis AI, Weiss A, Washington MK, et al. Heme sensing and detoxification by HatRT contributes to pathogenesis during *Clostridium difficile* infection. PLoS Pathog. 2018;14(12):e1007486. Epub 2018/12/21. doi:

10.1371/journal.ppat.1007486. PubMed PMID: 30576368; PubMed Central PMCID: PMCPMC6303022.

66. Kreiswirth BN, Löfdahl S, Betley MJ, O'Reilly M, Schlievert PM, Bergdoll MS, et al. The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. Nature. 1983;305(5936):709-12. doi: 10.1038/305709a0. PubMed PMID: 6226876.

67. Duthie ES, Lorenz LL. Staphylococcal coagulase; mode of action and antigenicity. J Gen Microbiol. 1952;6(1-2):95-107. doi: 10.1099/00221287-6-1-2-95. PubMed PMID: 14927856.

68. Attia AS, Benson MA, Stauff DL, Torres VJ, Skaar EP. Membrane damage elicits an immunomodulatory program in *Staphylococcus aureus*. PLoS Pathog. 2010;6(3):e1000802. Epub 2010/03/12. doi: 10.1371/journal.ppat.1000802. PubMed PMID: 20300601; PubMed Central PMCID: PMCPMC2837406.

69. Kehl-Fie TE, Chitayat S, Hood MI, Damo S, Restrepo N, Garcia C, et al. Nutrient metal sequestration by calprotectin inhibits bacterial superoxide defense, enhancing neutrophil killing of *Staphylococcus aureus*. Cell Host Microbe. 2011;10(2):158-64. doi: 10.1016/j.chom.2011.07.004. PubMed PMID: 21843872; PubMed Central PMCID: PMCPMC3157011.

70. Choby JE, Grunenwald CM, Celis AI, Gerdes SY, DuBois JL, Skaar EP. HemX Modulates Glutamyl-tRNA Reductase Abundance To Regulate Heme Biosynthesis. MBio. 2018;9(1). Epub 2018/02/06. doi: 10.1128/mBio.02287-17. PubMed PMID: 29437922; PubMed Central PMCID: PMCPMC5801465.

71. Schneewind O, Model P, Fischetti VA. Sorting of protein A to the staphylococcal cell wall. Cell. 1992;70(2):267-81. doi: 10.1016/0092-8674(92)90101-h. PubMed PMID: 1638631.

72. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, et al. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res. 2017;45(D1):D535-D42. Epub 2016/11/29. doi: 10.1093/nar/gkw1017. PubMed PMID: 27899627; PubMed Central PMCID: PMCPMC5210524.

73. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256-W9. doi: 10.1093/nar/gkz239. PubMed PMID: 30931475; PubMed Central PMCID: PMCPMC6602468.

74. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014;42(Database issue):D206-14. Epub 2013/11/29. doi: 10.1093/nar/gkt1226. PubMed PMID: 24293654; PubMed Central PMCID: PMCPMC3965101.

75. Seekatz AM, Theriot CM, Molloy CT, Wozniak KL, Bergin IL, Young VB. Fecal Microbiota Transplantation Eliminates *Clostridium difficile* in a Murine Model of Relapsing Disease. Infect Immun. 2015;83(10):3838-46. Epub 2015/07/13. doi: 10.1128/IAI.00459-15. PubMed PMID: 26169276; PubMed Central PMCID: PMCPMC4567621.

76. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, et al. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 2005;33(17):5691-702. Epub 2005/10/07. doi: 10.1093/nar/gki866. PubMed PMID: 16214803; PubMed Central PMCID: PMCPMC1251668.

77. Grove A. MarR family transcription factors. Curr Biol. 2013;23(4):R142-3. doi: 10.1016/j.cub.2013.01.013. PubMed PMID: 23428319.

78. Bérczi A, Zimányi L. The trans-membrane cytochrome b561 proteins: structural information and biological function. Curr Protein Pept Sci. 2014;15(8):745-60. PubMed PMID: 25163754.

79. Asard H, Barbaro R, Trost P, Bérczi A. Cytochromes b561: ascorbate-mediated trans-membrane electron transport. Antioxid Redox Signal. 2013;19(9):1026-35. doi: 10.1089/ars.2012.5065. PubMed PMID: 23249217; PubMed Central PMCID: PMCPMC3763232.

80. Wang Y, Branicky R, Noë A, Hekimi S. Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. J Cell Biol. 2018;217(6):1915-28. Epub 2018/04/18. doi: 10.1083/jcb.201708007. PubMed PMID: 29669742; PubMed Central PMCID: PMCPMC5987716.

81. Wu Y, Vulić M, Keren I, Lewis K. Role of oxidative stress in persister tolerance. Antimicrob Agents Chemother. 2012;56(9):4922-6. Epub 2012/07/09. doi: 10.1128/AAC.00921-12. PubMed PMID: 22777047; PubMed Central PMCID: PMCPMC3421885.

82. Karavolos MH, Horsburgh MJ, Ingham E, Foster SJ. Role and regulation of the superoxide dismutases of *Staphylococcus aureus*. Microbiology. 2003;149(Pt 10):2749-58. doi: 10.1099/mic.0.26353-0. PubMed PMID: 14523108.

83. Painter KL, Strange E, Parkhill J, Bamford KB, Armstrong-James D, Edwards AM. *Staphylococcus aureus* adapts to oxidative stress by producing H2O2-resistant small-colony variants via the SOS response. Infect Immun. 2015;83(5):1830-44. Epub 2015/02/17. doi: 10.1128/IAI.03016-14. PubMed PMID: 25690100; PubMed Central PMCID: PMCPMC4399076.

84. Van Acker H, Coenye T. The Role of Reactive Oxygen Species in Antibiotic-Mediated Killing of Bacteria. Trends Microbiol. 2017;25(6):456-66. Epub 2017/01/12. doi: 10.1016/j.tim.2016.12.008. PubMed PMID: 28089288.

85. Edwards DI. Nitroimidazole drugs--action and resistance mechanisms. II. Mechanisms of resistance. J Antimicrob Chemother. 1993;31(2):201-10. doi: 10.1093/jac/31.2.201. PubMed PMID: 8463167.

86. Kuehne SA, Collery MM, Kelly ML, Cartman ST, Cockayne A, Minton NP. Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. J Infect Dis. 2014;209(1):83-6. doi: 10.1093/infdis/jit426. PubMed PMID: 23935202; PubMed Central PMCID: PMCPMC3864386.

87. Rivera-Chávez F, Lopez CA, Bäumler AJ. Oxygen as a driver of gut dysbiosis. Free Radic Biol Med. 2017;105:93-101. Epub 2016/09/24. doi: 10.1016/j.freeradbiomed.2016.09.022. PubMed PMID: 27677568.

88. Kang JD, Myers CJ, Harris SC, Kakiyama G, Lee IK, Yun BS, et al. Bile Acid 7α-Dehydroxylating Gut Bacteria Secrete Antibiotics that Inhibit *Clostridium difficile*: Role of Secondary Bile Acids. Cell Chem Biol. 2019;26(1):27-34.e4. Epub 2018/10/25. doi: 10.1016/j.chembiol.2018.10.003. PubMed PMID: 30482679; PubMed Central PMCID: PMCPMC6338514.

89. Hill DR, Huang S, Nagy MS, Yadagiri VK, Fields C, Mukherjee D, et al. Bacterial colonization stimulates a complex physiological response in the immature human intestinal epithelium. Elife. 2017;6. Epub 2017/11/07. doi: 10.7554/eLife.29132. PubMed PMID: 29110754; PubMed Central PMCID: PMCPMC5711377.

90. Lopatkin AJ, Stokes JM, Zheng EJ, Yang JH, Takahashi MK, You L, et al. Bacterial metabolic state more accurately predicts antibiotic lethality than growth rate. Nat Microbiol. 2019. Epub 2019/08/26. doi: 10.1038/s41564-019-0536-0. PubMed PMID: 31451773.

91. Pericone CD, Park S, Imlay JA, Weiser JN. Factors contributing to hydrogen peroxide resistance in *Streptococcus pneumoniae* include pyruvate oxidase (SpxB) and avoidance of the toxic effects of the fenton reaction. J Bacteriol. 2003;185(23):6815-25. doi: 10.1128/jb.185.23.6815-6825.2003. PubMed PMID: 14617646; PubMed Central PMCID: PMCPMC262707.

92. Suits MD, Pal GP, Nakatsu K, Matte A, Cygler M, Jia Z. Identification of an *Escherichia coli* O157:H7 heme oxygenase with tandem functional repeats. Proc Natl Acad Sci U S A. 2005;102(47):16955-60. Epub 2005/11/07. doi: 10.1073/pnas.0504289102. PubMed PMID: 16275907; PubMed Central PMCID: PMCPMC1287972.

93. Stojiljkovic I, Hantke K. Transport of haemin across the cytoplasmic membrane through a haemin-specific periplasmic binding-protein-dependent transport system in *Yersinia enterocolitica*. Mol Microbiol. 1994;13(4):719-32. PubMed PMID: 7997183.

94. Marvig RL, Damkiær S, Khademi SM, Markussen TM, Molin S, Jelsbak L. Within-host evolution of *Pseudomonas aeruginosa* reveals adaptation toward iron acquisition from hemoglobin. MBio. 2014;5(3):e00966-14. Epub 2014/05/06. doi: 10.1128/mBio.00966-14. PubMed PMID: 24803516; PubMed Central PMCID: PMCPMC4010824.

95. Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. Microbiol Mol Biol Rev. 1998;62(1):1-34. PubMed PMID: 9529885; PubMed Central PMCID: PMCPMC98904.

96. Edwards AN, Karim ST, Pascual RA, Jowhar LM, Anderson SE, McBride SM. Chemical and Stress Resistances of *Clostridium difficile*. Front Microbiol. 2016;7:1698. Epub 2016/10/26.

doi: 10.3389/fmicb.2016.01698. PubMed PMID: 27833595; PubMed Central PMCID: PMCPMC5080291.

APPENDIX A

Tables associated with Chapter II

Appendix A Table 1 *C. difficile* R20291 genes transcriptionally upregulated in the presence of heme.

Gene Symbol	Fold-change Heme Treated vs Untreated Media Control	Description
CDR20291_0781	31.78261	putative membrane protein
CDR20291_0782	26.463398	MarR-family transcriptional regulator
CDR20291_1227 (*hatR)*	11.18395	TetR-family transcriptional regulator
CDR20291_2306	6.155928	conserved hypothetical protein
CDR20291_1226 (*hatT)*	5.054025	putative transporter
tcdD	4.6314363	putative transcriptional regulator
ctfA	4.446509	butyrate-acetoacetate CoA-transferase subunit A
CDR20291_0433	4.170941	putative sugar-phosphate dehydrogenase
CDR20291_2321	3.9144623	putative oxidoreductase ferredoxin subunit
CDR20291_1223	3.6944113	putative phage regulatory protein
oraS	3.594828	D-ornithine aminomutase S component
thlA2	3.492519	putative acetyl-CoA acetyltransferase
CDR20291_2801	3.4535594	putative membrane protein
CDR20291_0163	3.4244125	hypothetical protein
CDR20291_0284	3.4145832	conserved hypothetical protein
CDR20291_3277	3.3464322	putative exported protein
CDR20291_1726	3.3368134	hypothetical protein
CDR20291_0391	3.3224375	putative amino acid racemase
acpP	3.1869857	acyl carrier protein
CDR20291_1554	3.136697	putative membrane protein
CDR20291_3464	3.0761995	conjugative transposon protein
CDR20291_1445	3.0670178	hypothetical phage protein
CDR20291_0033	3.0560584	putative membrane protein
CDR20291_2677	3.034006	putative cell wall teichoic acid glycosylation protein
CDR20291_0347	3.0311313	hypothetical protein
spoIIIAG	3.0204601	stage iii sporulation protein ag flags: precursor
CDR20291_3191	3.0179818	conserved hypothetical protein
CDR20291_0738	2.9525864	hypothetical protein
CDR20291_0739	2.9086928	putative membrane protein

CDR20291_1810	2.8909168	hypothetical protein
CDR20291_0291	2.8904557	PTS system, IIc component
CDR20291_2194	2.8736017	putative membrane protein precursor
CDR20291_0141	2.850998	putative RNA-binding protein
potD	2.8405726	spermidine/putrescine ABC transporter, substrate-binding lipoprotein
CDR20291_2335	2.8380623	conserved hypothetical protein
CDR20291_2575	2.8209536	hypothetical protein
u1	2.8027954	putative regulatory protein
CDR20291_1667	2.77625	hypothetical protein
CDR20291_1393	2.7436821	hypothetical protein
CDR20291_0398	2.7361345	putative membrane protein
CDR20291_3511	2.7244775	putative peptidase
CDR20291_0728	2.7207413	putative hydroxymethylglutaryl-CoA lyase
CDR20291_0214	2.7207377	putative nitroreductase
CDR20291_2039	2.7174134	conserved hypothetical protein
CDR20291_2727	2.6998487	putative peptidase
CDR20291_2190	2.6926663	putative regulatory protein
CDR20291_2517	2.6859193	putative transcriptional regulator
CDR20291_3465	2.6800773	conjugative transposon protein
CDR20291_1866	2.6749303	conserved hypothetical protein
CDR20291_1658	2.6748435	putative membrane protein
CDR20291_0390	2.6627486	putative component of D-ornithine aminomutase
CDR20291_3094	2.6538692	hypothetical protein
CDR20291_2698	2.6518354	conserved hypothetical protein
CDR20291_1823	2.6514072	putative lipoprotein signal peptidase
CDR20291_3453	2.649938	putative collagen-binding surface protein
CDR20291_1078	2.6334527	hypothetical protein
CDR20291_3460	2.6278205	conjugative transposon protein
CDR20291_1073	2.6088905	hypothetical protein
CDR20291_2728	2.6057155	putative membrane protein
CDR20291_3466	2.5999343	putative cell wall hydrolase
CDR20291_0587	2.592949	hypothetical protein
pgmB	2.5752125	beta-phosphoglucomutase
CDR20291_2189	2.5734594	putative repressor
ydiB	2.5728512	NAD-dependent shikimate 5-dehydrogenase

CDR20291_0288	2.5664334	PTS system, IIb component
CDR20291_3140	2.555809	PTS system, IIb component
CDR20291_1971	2.5507545	hypothetical protein
CDR20291_0294	2.5414681	putative peptidase
rpmH	2.5404537	50S ribosomal protein L34
CDR20291_1083	2.5396543	putative membrane protein
CDR20291_3153	2.529364	putative membrane protein
CDR20291_1319	2.5218627	putative phage shock protein
CDR20291_3463	2.5192652	conjugative tranposon protein
fdxA	2.517662	ferredoxin
CDR20291_3132	2.504153	hypothetical protein
CDR20291_0409	2.501609	putative hydrolase
CDR20291_1688	2.4785788	putative membrane protein
CDR20291_1026	2.4770615	putative glutamine amidotransferase
CDR20291_2775	2.476111	hypothetical protein
CDR20291_1820	2.4603844	hypothetical protein
CDR20291_3053	2.4462643	putative phage-related protein
rpsO	2.431163	30S ribosomal protein S15
CDR20291_1192	2.419036	putative lantibiotic ABC transporter, permease protein
CDR20291_0203	2.4166424	conserved hypothetical protein
CDR20291_3113	2.415126	two-component response regulator
CDR20291_2019	2.412923	hypothetical protein
CDR20291_1222	2.409677	putative phage regulatory protein
CDR20291_1430	2.4094286	hypothetical protein
CDR20291_1863	2.4049656	putative membrane protein
CDR20291_1156	2.4041753	hypothetical protein
spoVG	2.4039543	stage V sporulation protein G
CDR20291_1618	2.4028168	conserved hypothetical protein
CDR20291_2405	2.4002483	putative translation inhibitor endoribonuclease
CDR20291_1576	2.3943622	hypothetical protein (pseudogene)
secG	2.393486	putative subunit of preprotein translocase
CDR20291_3505	2.3905718	GntR-family transcriptional regulator
CDR20291_2807	2.3885088	hypothetical protein
phnA	2.3848896	putative phosphonoacetate hydrolase
CDR20291_2953	2.384776	putative membrane protein
CDR20291_1949	2.3762958	putative uncharacterized protein

CDR20291_3187A	2.3746626	autoinducer prepeptide
CDR20291_2396	2.372154	putative D-alanyl-D-alanine carboxypeptidase
sat	2.36493	streptogramin A acetyltransferase
rpmE	2.359248	50S ribosomal protein L31
CDR20291_1126	2.3415961	conserved hypothetical protein
CDR20291_3110	2.3412015	conserved hypothetical protein
CDR20291_3387	2.336615	conserved hypothetical protein
CDR20291_3458	2.3194883	putative conjugative transposon FtsK_SpoIIIE-related protein
CDR20291_2106	2.315871	putative oxidoreductase, ferredoxin subunit
cspA	2.3125906	cold shock protein
CDR20291_1336	2.3109813	hypothetical protein
CDR20291_2709	2.3093321	transposase
CDR20291_3285	2.3065357	putative uncharacterized protein flags: precursor
CDR20291_3108	2.304577	conserved hypothetical protein
CDR20291_1672	2.2980769	putative arsenate reductase
CDR20291_3448	2.293609	hypothetical protein
CDR20291_2916	2.2925122	phosphosugar-binding transcriptional regulator
CDR20291_1191	2.2901857	putative lantibiotic ABC transporter, ATP-binding protein
CDR20291_0418	2.289149	hypothetical protein
CDR20291_2712	2.2884893	putative peptidase
CDR20291_2361	2.2881973	putative exported protein
CDR20291_0155	2.286317	putative membrane-associated CAAX amino terminal protease
d1	2.2808263	conserved hypothetical protein
CDR20291_3293	2.2803311	putative exported protein
CDR20291_1229	2.2741683	GntR-family transcriptional regulator
CDR20291_2375	2.2706938	putative lipoprotein
CDR20291_3462	2.2694504	conjugative transposon protein
CDR20291_2977	2.2647276	transcription antiterminator
CDR20291_0665	2.2631786	putative exported protein
oraE	2.2565958	D-ornithine aminomutase E component
CDR20291_0511	2.253925	conserved hypothetical protein
CDR20291_0775	2.2506297	putative nuclease
glvC	2.2502027	PTS system, IIbc component
CDR20291_3400	2.2420852	putative spore cortex-lytic enzyme

CDR20291_1928	2.2351036	putative membrane protein
CDR20291_1176	2.2335684	MarR-family transcriptional regulator
CDR20291_1481	2.2287	conserved hypothetical protein
CDR20291_2499	2.2240226	hypothetical protein
CDR20291_2651	2.2223415	putative lipoprotein
CDR20291_3467	2.2171688	conjugative transposon protein
sspB	2.2168262	small acid-soluble spore protein B
CDR20291_1142	2.2137163	hypothetical protein
CDR20291_3120	2.21356	putative phosphateABC transporter, permease protein
spoIIIAA	2.2119963	stage III sporulation protein AA
CDR20291_2808	2.2068152	conserved hypothetical protein (fragment)
CDR20291_3504	2.204511	PTS system, IIc component
CDR20291_0543	2.2003973	conserved hypothetical protein
CDR20291_1461	2.19731	holin
CDR20291_3417	2.193156	conserved hypothetical protein
glvR	2.1895297	phosphosugar-binding transcriptional regulator
CDR20291_1527	2.1887293	putative membrane protein
CDR20291_0956	2.1862566	hypothetical protein
srlB	2.179912	PTS system, IIa component
CDR20291_1707	2.1784844	putative two-component system response regulator
fur	2.1756976	ferric uptake regulation protein
licT	2.1745243	putative transcription antiterminator
CDR20291_2490	2.1732683	putative response regulator
CDR20291_0560	2.1706605	hypothetical protein
CDR20291_0697	2.1640105	transposase-like protein b
CDR20291_1639	2.1629908	putative ferrous iron transport protein A
CDR20291_2087	2.1621017	putative aromatic compounds hydrolase
CDR20291_1559	2.1569982	conserved hypothetical protein
CDR20291_1914	2.1558685	hypothetical protein
rbsA	2.1556127	ribose ABC transporter, ATP-binding protein
CDR20291_2917	2.1528287	putative membrane protein
CDR20291_3064	2.1512	putative membrane protein
CDR20291_2182	2.1462066	hypothetical protein
CDR20291_0512	2.144096	hypothetical protein

CDR20291_3009	2.1413686	hypothetical protein
CDR20291_1225	2.1394012	putative phage regulatory protein
CDR20291_0859	2.1387832	conserved hypothetical protein
CDR20291_1447	2.1384666	putative uncharacterized protein
CDR20291_0796	2.1379225	putative membrane protein
CDR20291_2281	2.1370592	putative membrane protein precursor
CDR20291_0571	2.13593	putative peptidase
CDR20291_2320	2.1343033	putative oxidoreductase subunit
spoIIR	2.1340494	stage II sporulation protein
CDR20291_1323	2.133574	putative ruberythrin
CDR20291_3106	2.1280892	probable polysaccharide deacetylase
CDR20291_1496	2.1274724	putative endonuclease
CDR20291_3461	2.1273339	chloramphenicol o-acetyltransferase
CDR20291_3408	2.1271412	hypothetical protein
CDR20291_0207	2.123493	PTS system, IIa component
CDR20291_2979	2.1226456	putative sugar-bisphosphate aldolase
CDR20291_2286	2.122154	conserved hypothetical protein
CDR20291_0340	2.1204524	hypothetical protein
CDR20291_1610	2.120411	conserved hypothetical protein
CDR20291_2336	2.120107	putative sigma 54 modulation protein
veg	2.1170819	conserved hypothetical protein
CDR20291_3154	2.113426	hypothetical protein
CDR20291_3444	2.1119018	ferredoxin
adhE	2.1084027	aldehyde-alcohol dehydrogenase
srlA	2.1078258	PTS system, glucitol/sorbitol-specific IIc2 component
CDR20291_2939	2.1050506	PTS system, IIbc component pts system, iibc component precursor
CDR20291_0209	2.0995953	PTS system, IIb component
CDR20291_0614	2.099349	conserved hypothetical protein
CDR20291_0029	2.0985513	putative transcription antiterminator
CDR20291_1738	2.0930097	conserved hypothetical protein
CDR20291_2649	2.0913396	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_2192	2.0892425	transposase (fragment)
CDR20291_0184	2.086086	putative cell wall hydrolase
CDR20291_2256	2.0859647	conserved hypothetical protein

CDR20291_0158	2.0842288	putative two-component response regulator
CDR20291_3350	2.081579	pilin
CDR20291_2952	2.0789118	putative amidohydrolase
CDR20291_2690	2.0775318	putative protein translocase subunit
CDR20291_2856	2.0749261	conserved hypothetical protein
rbsC	2.070177	ribose ABC transporter, permease protein
CDR20291_0942	2.06901	PTS system, IIb component
CDR20291_1813	2.0665524	putative regulatory protein
CDR20291_2285	2.0649927	putative membrane protein
CDR20291_2066	2.0603025	putative membrane protein
CDR20291_2233	2.058511	putative membrane protein
CDR20291_0344	2.0582306	two-component response regulator
CDR20291_0760	2.0540824	putative membrane protein
CDR20291_3214	2.049898	hypothetical protein
CDR20291_0928	2.048106	putative phosphatTse
CDR20291_0287	2.0476503	PTS system, IIa component
CDR20291_2372	2.0469162	conserved hypothetical protein
CDR20291_0022	2.044626	putative beta-xylosidase
CDR20291_3498	2.0391083	LysR-family transcriptional regulator
abgT	2.0381641	putative aminobenzoyl-glutamate transport protein
CDR20291_2498	2.0380943	TetR-family trancscriptional regulator
eutN	2.0369942	putative ethanolamine/propanediol utilization protein
CDR20291_0171	2.0355306	putative redox-sensing transcriptional repressor
CDR20291_1258	2.033891	conserved hypothetical protein
CDR20291_1025	2.033765	conserved hypothetical protein
CDR20291_3356	2.0323858	putative exported protein
CDR20291_0514	2.0303247	transposase-like protein b pseudogene
CDR20291_2868	2.0281112	thioredoxin
CDR20291_0875	2.0269256	conserved hypothetical protein
CDR20291_1619	2.0243516	putative transcriptional regulator
CDR20291_0021	2.0238392	putative beta-glucosidase
CDR20291_1138	2.021514	conserved hypothetical protein
sleB	2.0209548	putative spore-cortex-lytic protein

CDR20291_0133	2.0206575	putative transcription antiterminator
cspD	2.0195987	cold shock protein
rpoD2	2.0195546	RNA polymerase sigma factor rpoD
CDR20291_1472	2.0193841	putative exported protein
CDR20291_3480	2.0188391	conserved hypothetical protein
spoIIIAC	2.0182188	stage III sporulation protein AC
spoVT	2.015846	stage V sporulation protein T
CDR20291_0385	2.0120325	putative oxidoreductase
CDR20291_2798	2.0021243	conserved hypothetical protein
CDR20291_2545	2.0008383	putative membrane protein
CDR20291_1241	2.0001361	putative membrane protein

Appendix A Table 2 *C. difficile* R20291 genes transcriptional downregulated in the presence of heme.

Gene Symbol	Fold-change Heme Treated vs Untreated Media Control	Description
cotJB1	-8.493219	putative spore-coat protein
CDR20291_0521	-5.746132	hypothetical protein
cotJC1	-3.9827776	putative spore-coat protein
rpsE	-3.8330042	30S ribosomal protein S5
rplV	-3.7833204	50S ribosomal protein L22
CDR20291_3239	-3.7300813	ABC transporter, ATP-binding protein
CDR20291_3046	-3.705808	MerR-family transcriptional regulator
rplP	-3.7040238	50S ribosomal protein L16
rpsH	-3.6383638	30S ribosomal protein S8
rpsS	-3.6113067	30S ribosomal protein S19
rplR	-3.604462	50S ribosomal protein L18
rpsQ	-3.5911934	30S ribosomal protein S17
rpsN	-3.5626237	30S ribosomal protein S14
rplN	-3.5618937	50S ribosomal protein L14
cotJB2	-3.5552454	putative spore-coat protein
feoB1	-3.4932115	ferrous iron transport protein B
CDR20291_1075	-3.434818	putative exported protein
rplX	-3.4215739	50S ribosomal protein L24
rpsC	-3.3353822	30S ribosomal protein S3
rplE	-3.283783	50S ribosomal protein L5
rplD	-3.1962414	50S ribosomal protein L4
rpmD	-3.1648352	50S ribosomal protein L30
rplW	-3.136498	50S ribosomal protein L23
rplF	-3.1316626	50S ribosomal protein L6
rpmC	-3.102201	50S ribosomal protein L29
rplB	-3.0742095	50S ribosomal protein L2
CDR20291_3237	-3.0330143	ABC transporter, permease protein
rplC	-3.0299876	50S ribosomal protein L3
CDR20291_1648	-2.9830742	putative ABC transporter, permease protein
fhuG	-2.8928006	putative ferrichrome ABC transporter, permease protein
CDR20291_0926	-2.8338807	hypothetical protein
adk	-2.8257287	adenylate kinase
rpsK	-2.818008	30S ribosomal protein S11
CDR20291_1150	-2.7863884	putative ribosomal protein
CDR20291_1210	-2.765053	putative phage protein

CDR20291_1149	-2.7456932	conserved hypothetical protein
ribA	-2.7125893	riboflavin biosynthesis protein
CDR20291_3099	-2.6862726	conserved hypothetical protein
infB	-2.6825497	translation initiation factor IF-2
sodA	-2.6655688	putative superoxide dismutase [Mn]
CDR20291_0922	-2.635544	hypothetical protein
tuaG	-2.6347167	putative teichuronic acid biosynthesis glycosyl transferase
CDR20291_0516	-2.6060565	putative cation transporting ATPase
CDR20291_3238	-2.6003761	ABC transporter, permease protein
CDR20291_1212	-2.585924	putative phage cell wall hydrolase
CDR20291_2449	-2.5827234	conserved hypothetical protein
CDR20291_0996	-2.5810926	conserved hypothetical protein
fhuC	-2.5564084	putative ferrichrome ABC transporter, ATP-binding protein
dapA2	-2.5528688	dihydrodipicolinate synthase
fhuB	-2.5447285	putative ferrichrome ABC transporter, permease protein
CDR20291_1647	-2.542131	ABC transporter, ATP-binding protein
sigH	-2.5233936	RNA polymerase sigma-H factor
CDR20291_3100	-2.5013387	conserved hypothetical protein
CDR20291_2138	-2.492333	putative sodium:solute symporter
flgG	-2.4588897	flagellar basal-body rod protein FlgG
CDR20291_2289	-2.437157	hypothetical protein
aroC	-2.420682	chorismate synthase (5-enolpyruvylshikimate-3-phosphate phospholyase)
rpsM	-2.4126081	30S ribosomal protein S13
CDR20291_3022	-2.409462	conserved hypothetical protein
CDR20291_1329	-2.384787	putative exported protein
CDR20291_1538	-2.3766038	conserved hypothetical protein
CDR20291_1581	-2.371138	putative membrane protein
thiK	-2.3697405	4-methyl-5-beta-hydroxyethylthiazole kinase hydroxyethylthiazole kinase
CDR20291_1214	-2.3692572	phage protein
CDR20291_1649	-2.3312683	putative ABC transporter, permease protein
CDR20291_1216	-2.329888	phage protein
CDR20291_2661	-2.3286161	putative beta-glycosyltransferase
CDR20291_1211	-2.3246927	putative phage protein
CDR20291_0346	-2.3085475	conserved hypothetical protein
metN	-2.2932603	D-methionine ABC transporter, ATP-binding protein

CDR20291_2770	-2.2888985	putative drug/sodium antiporter
rpsD	-2.2606547	30S ribosomal protein S4
flgE	-2.2575037	flagellar hook protein
fchA	-2.246807	methenyltetrahydrofolate cyclohydrolase
obg	-2.246415	Spo0B-associated GTP-binding protein
csrA	-2.2385561	carbon storage regulator
rpmJ	-2.233	50S ribosomal protein L36
CDR20291_3006	-2.2250257	putative phage protein
dapB1	-2.2225766	dihydrodipicolinate reductase
CDR20291_0746	-2.2034466	probable transporter
CDR20291_0043	-2.197775	thymidylate synthase
fabG	-2.1970513	3-oxoacyl-[acyl-carrier protein] reductase
gutA	-2.196997	PTS system, glucitol/sorbitol-specific IIc2 component
CDR20291_0273	-2.1895223	putative flagellar basal-body rod protein
flgL	-2.1887264	flagellar hook-associated protein
acoB	-2.1874819	acetoin:2,6-dichlorophenolindophenol oxidoreductase beta subunit
CDR20291_0228	-2.1758933	conserved hypothetical protein
asd	-2.1721008	aspartate-semialdehyde dehydrogenase
CDR20291_2814	-2.1702242	conserved hypothetical protein
rpsJ	-2.16123	30S ribosomal protein S10
hisS	-2.1584945	putative histidyl-tRNA synthetase
CDR20291_0045	-2.146511	putative uncharacterized protein
CDR20291_0047	-2.1418478	putative thymidylate synthase
CDR20291_2662	-2.1404283	putative teichuronic acid biosynthesis glycosyl transferase
CDR20291_1217	-2.1344755	putative phage tail fiber protein
uvrB	-2.1312566	excinuclease ABC subunit B
iunH	-2.1290586	inosine-uridine preferring nucleoside hydrolase
CDR20291_1551	-2.1245148	putative lipoprotein
rplO	-2.1228526	50S ribosomal protein L15
CDR20291_1125	-2.1123612	putative holliday junction resolvase
CDR20291_0804	-2.1063793	ABC transporter, permease protein
feoA1	-2.1019013	putative ferrous iron transport protein A
cheC	-2.0987973	chemotaxis protein CheC
sip2	-2.0981913	singal peptidase I
nusA	-2.0957475	transcription elongation protein
CDR20291_0995	-2.0921319	radical SAM-superfamily protein

hadB	-2.0901814	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
tpi	-2.0859563	triosephosphate isomerase
CDR20291_0648	-2.0853581	conserved hypothetical protein
CDR20291_1218	-2.085296	putative phage protein
CDR20291_3102	-2.0845835	hypothetical protein
rnfE	-2.0844994	electron transport complex protein
rnfA	-2.0843024	electron transport complex protein
fleN	-2.0812418	flagellar number regulator
CDR20291_2884	-2.0809321	putative PTS system, IIb component
CDR20291_2346	-2.0776253	conserved hypothetical protein
folD	-2.0760767	putative FolD bifunctional protein
prlA	-2.0753083	preprotein translocase SecY subunit
CDR20291_1213	-2.0732968	hypothetical protein
CDR20291_0065	-2.0723321	elongation factor TU
CDR20291_0985	-2.0693574	putative penicillin-binding protein
CDR20291_0044	-2.061752	dihydrofolate reductase region
fliS1	-2.0603256	flagellar protein FliS
CDR20291_2349	-2.0602372	ABC transporter, ATP-binding protein
manC	-2.0563254	putative mannose-1-phosphate guanylyltransferase
CDR20291_2526	-2.0555916	two-component response regulator
CDR20291_0049	-2.0538368	conserved hypothetical protein
pgm2	-2.0533345	putative phosphomannomutase/phosphoglycerate mutase
CDR20291_1103	-2.0531716	putative FMN-dependent dehydrogenase
map1	-2.0526085	methionine aminopeptidase
CDR20291_2825	-2.0481207	ABC transporter, ATP-binding protein
CDR20291_0227	-2.0478148	putative transglycosylase
rpiB1	-2.0455568	ribose-5-phosphate isomerase 1
cheD	-2.0426855	chemotaxis protein
hadI	-2.034394	activator of 2-hydroxyisocaproyl-CoA dehydratase
mreC	-2.029631	putative rod shape-determining protein precursor
CDR20291_1615	-2.0271087	probable permease
pheT	-2.023367	phenylalanyl-tRNA synthetase beta chain
CDR20291_1107	-2.0229175	putative ABC transporter, permease protein
thiI	-2.0227945	putative thiamine biosynthesis protein

CDR20291_2911	-2.0177546	restriction modification system dna specificity domain
CDR20291_0463	-2.0165677	putative methyl accpeting chemotaxis protein
feoB3	-2.0153618	putative ferrous iron transport protein B
aspS	-2.010659	putative aspartyl-tRNA synthetase
rpe	-2.0086982	putative ribulose-phosphate 3-epimerase
flgD	-2.00112	putative basal-body rod modification protein

APPENDIX B

Tables associated with Chapter III

Appendix B Table 1 *C. difficile* R20291 genes in the presence of heme (fold change greater or lesser than 2)

Gene Symbol	Fold-change Heme Treated vs Untreated Media Control	Description
CDR20291_0782 (*hsmR)*	6.64856	putative membrane protein
CDR20291_0781 (*hsmA)*	5.63898	MarR-family transcriptional regulator
hatR	4.35365	heme sensing transcriptional regulator
hatT	4.20734	heme detoxification transporter
hcp	-2.02141	hydroxylamine reductase

Appendix B Table 2 *C. difficile hsmR::CT* genes in the presence of heme (fold change greater or lesser than 2)

Gene Symbol	Fold-change Heme Treated vs Untreated Media Control	Description
hatR	5.767794	heme sensing transcriptional regulator
hatT	5.436752	heme detoxification transporter

Appendix Table 3 *C. difficile* WT and *hsmR::CT* genes in the absence of heme (fold change greater or lesser than 2)

Gene Symbol	Fold-change WT vs *hsmR::CT*	Description
CDR20291_2480	2.836956	macrolide transport system ATP-binding/permease protein
CDR20291_0782 (*hsmR)*	2.52901	MarR-family transcriptional regulator
hatR	4.35365	heme sensing transcriptional regulator
hatT	4.20734	heme detoxification transporter
hcp	-2.02141	hydroxylamine reductase
CDR20291_0226	-2.00855	dDTP-glucose 4,6-dehydratase
flgD	-2.02181	basal-body rod modification protein
CDR20291_0225	-2.03484	dDTP-4-dehydrorhamnose reductase
motA	-2.05794	chemotaxis protein
CDR20291_0224	-2.06797	glucose-1-phosphate thymidylyltransferase
fliQ	-2.10368	flagellar biosynthetic protein
fliG	-2.15122	Flagellar motor switch protein
CDR20291_0227	-2.16968	putative transglycosylase
fliF	-2.18226	M-ring protein
CDR20291_0228	-2.19314	conserved hypothetical protein
CDR20291_0241	-2.23863	putative glycosyltransferase
fliJ	-2.23974	flagellar protein
fliL	-2.26158	flagellar basal body-associated protein
flgB	-2.26292	basal-body rod protein
CDR20291_0263	-2.33069	putative flagellar protein
flgC	-2.3508	flagellar basal-body rod protein
CDR20291_0259	-2.35087	putative flagellar protein
fliN	-2.3789	flagellar motor switch protein
fliE	-2.40667	hook-basal body complex protein
flgL	-2.41864	flagellar hook-associated protein
fliD	-2.44488	flagellar cap protein
flgK	-2.45269	flagellar hook-associated protein
fliS1	-2.4758	flagellar protein
flgM	-2.49579	negative regulator of flagellin synthesis
fliS2	-2.52948	flagellar protein
CDR20291_0239	-2.57365	conserved hypothetical protein
csrA	-2.57683	carbon storage regulator
CDR20291_0234	-2.61381	conserved hypothetical protein
CDR20291_0231	-2.63196	putative flagellar biosynthesis protein
fliC	-2.66162	flagellin subunit

Appendix B Table 4 *C. difficile* WT and *hsmR::CT* genes in the presence of heme (fold change greater or lesser than 2)

Gene Symbol	Fold-change WT vs *hsmR::CT*	Description
CDR20291_2480	2.444228	macrolide transport system ATP-binding/permease protein
CDR20291_0782 (*hsmR)*	2.52901	MarR-family transcriptional regulator
CDR20291_0226	-2.00855	dDTP-glucose 4,6-dehydratase
flhF	-2.00144	flagellar biosynthesis protein
CDR20291_0245	-2.01568	putative carbamoyl-phosphate-synthetase
fliL	-2.03461	flagellar basal body-associated protein
CDR20291_0273	-2.07945	putative flagellar basal-rod protein
fliA	-2.07994	RNA polymerase sigma factor for flagellar operon
flgE	-2.08606	flagellar hook protein
CDR20291_0228	-2.09306	conserved hypothetical protein
fliJ	-2.10666	flagellar protein
CDR20291_0242	-2.10746	glucosyl transferase
CDR20291_0243	-2.10907	glucosyl transferase
fliI	-2.12922	flagellar basal body-associated protein
motA	-2.14247	chemotaxis protein
CDR20291_0224	-2.14944	glucose-1-phosphate thymidylyltransferase
fliG	-2.16719	Flagellar motor switch protein
CDR20291_0263	-2.19204	putative flagellar protein
flgC	-2.19752	flagellar basal-body rod protein
motB	-2.24352	chemotaxis protein
flgD	-2.28087	basal-body rod modification protein
CDR20291_0226	-2.29491	dTDP-glucose 4,6-dehydratase
csrA	-2.3151	carbon storage regulator
fliH	-2.33304	flagellar assembly protein
CDR20291_0231	-2.33866	putative flagellar biosynthesis protein
fliF	-2.34363	M-ring protein
CDR20291_0225	-2.34945	dDTP-4-dehydrorhamnose reductase
fliK	-2.3646	hook-length control protein
fliC	-2.37099	flagellin subunit
fliE	-2.42845	hook-basal body complex protein
flgB	-2.44719	basal-body rod protein

fliS2	-2.4692	flagellar protein
flgM	-2.50214	negative regulator of flagellin synthesis
fliD	-2.55771	flagellar cap protein
CDR20291_0234	-2.55999	conserved hypothetical protein
CDR20291_0241	-2.56217	putative glycosyltransferase
flgL	-2.57338	flagellar hook-associated protein
flgK	-2.58972	flagellar hook-associated protein
CDR20291_0239	-2.67364	conserved hypothetical protein
fliS1	-2.69584	flagellar protein
CDR20291_0782 (*hsmR*)	-4.30498	MarR-family transcriptional regulator
CDR20291_0781 (*hsmA*)	-5.47692	putative membrane protein

Appendix B Table 5 Bacterial species containing HsmA orthologs

Genome	Genome ID	Accession	PATRIC ID	RefSeq Locus Tag
Acetanaerobacterium elongatum strain CGMCC 1.5012	258515.18	FNID01000013	fig|258515.18.peg.672	SAMN05192585_1132
Acetobacterium sp. KB-1	2184575.3	CP030040	fig|2184575.3.peg.724	DOZ58_03405
Acetobacterium sp. MES1	1899015.3	MJUY01000013	fig|1899015.3.peg.924	BI182_02220
Acetobacterium sp. MES1	1899015.4	MJUY01000013	fig|1899015.4.peg.924	BI182_02220
Acetobacterium woodii DSM 1030	931626.3	NC_016894	fig|931626.3.peg.1649	Awo_c15450
Actinobacteria bacterium 66_15	1635289.5	LGFV01000013	fig|1635289.5.peg.191	XD74_0548
Actinobacteria bacterium GWC2_53_9	1797195.3	MELI01000051	fig|1797195.3.peg.712	A2074_02735
Actinobacteria bacterium HGW-Actinobacteria-10	2013645.3	PHFA01000014	fig|2013645.3.peg.65	CVT60_02245
Actinobacteria bacterium HGW-Actinobacteria-6	2013651.3	PHEU01000021	fig|2013651.3.peg.2201	CVT66_07775
Actinobacteria bacterium HGW-Actinobacteria-7	2013652.3	PHET01000006	fig|2013652.3.peg.473	CVT67_04560
Actinobacteria bacterium strain UBA10029	1883427.117	DPGQ01000212	fig|1883427.117.peg.829	DE036_09630
Actinobacteria bacterium strain UBA10799	1883427.106	DMNR01000232	fig|1883427.106.peg.714	DCQ04_08745
Actinomyces bovis strain NCTC11535	1658.3	UAPQ01000010	fig|1658.3.peg.2059	NCTC11535_01894
Actinomyces israelii strain NCTC12972	1659.4	LR134357	fig|1659.4.peg.2192	NCTC12972_02198
Actinomyces sp. VUL4_3	1912795.3	CP017812	fig|1912795.3.peg.353	BK816_01655
Agromyces sp. CF514	1881031.3	FOZD01000001	fig|1881031.3.peg.1499	SAMN05428970_0176
Anaerocolumna jejuensis DSM 15929	1121322.3	FRAC01000018	fig|1121322.3.peg.1163	SAMN02745136_03470
Anaerofustis stercorihominis DSM 17244	445971.6	NZ_DS560019	fig|445971.6.peg.1247	ANASTE_01335
Anaerofustis stercorihominis strain AM25-6	214853.3	QUSM01000002	fig|214853.3.peg.595	DW687_02905
Anaerolineaceae bacterium 4572_5.2 strain 4572_5.2	1971725.3	NBMA01000100	fig|1971725.3.peg.247	B6243_05835
A*naerolineaceae* bacterium strain NAT123	2024896.3	PABI01000076	fig|2024896.3.peg.2696	CL608_12265

Anaerolineae bacterium CG03_land_8_20_14_0_80_58_20	1973905.3	PEUL01000403	fig|1973905.3.peg.624	COS37_06875
Anaerolineae bacterium CG1_02_58_13	1805002.3	MNUH01000400	fig|1805002.3.peg.1716	AUJ21_11165
Anaerolineales bacterium strain FeB_25	2073117.4	PQAL01000013	fig|2073117.4.peg.1328	C3F13_06355
Andreesenia angusta strain DSM 1989	39480.4	MKIE01000009	fig|39480.4.peg.1967	EUAN_20010
Aneurinibacillus soli strain CB4	1500254.3	AP017312	fig|1500254.3.peg.2773	CB4_02761
Arthrobacter alpinus strain ERGS4:06	656366.4	CP013200	fig|656366.4.peg.2879	AS189_13375
Arthrobacter sp. AQ5-05	2184581.3	QMKQ01000021	fig|2184581.3.peg.3344	DQ353_15850
Arthrobacter sp. HMWF013	2056849.3	QAIQ01000534	fig|2056849.3.peg.4048	DBR22_18455
Atopobium sp. oral taxon 810 str. F0209	1321773.3	AWSK01000013	fig|1321773.3.peg.414	HMPREF9069_00469
Bacillus badius strain DSM 5610	1455.7	LVTO01000018	fig|1455.7.peg.2533	A3781_16875
Bacillus cereus HuA4-10	1053206.3	AHEA01000021	fig|1053206.3.peg.2575	IGC_02524
Bacillus cereus Rock3-44	526986.3	NZ_CM000733	fig|526986.3.peg.4201	bcere0022_7590
Bacillus cereus strain AFS002368	1396.1405	NTZF01000004	fig|1396.1405.peg.3340	CN491_04140
Bacillus cereus strain AFS005615	1396.1417	NTYD01000046	fig|1396.1417.peg.2880	CN476_21740
Bacillus cereus strain AFS010695	1396.1496	NTWE01000031	fig|1396.1496.peg.3243	CN425_17710
Bacillus cereus strain AFS016962	1396.1476	NTUE01000101	fig|1396.1476.peg.29	CN382_30750
Bacillus cereus strain AFS024089	1396.1455	NTRE01000040	fig|1396.1455.peg.3442	CN285_16210
Bacillus cereus strain AFS028441	1396.1430	NTSZ01000013	fig|1396.1430.peg.647	CN354_05365
Bacillus cereus strain AFS031783	1396.1828	NUOT01000008	fig|1396.1828.peg.5041	COE15_02105
Bacillus cereus strain AFS036423	1396.1805	NUNH01000058	fig|1396.1805.peg.4625	COD94_21615
Bacillus cereus strain AFS050027	1396.1758	NUIL01000056	fig|1396.1758.peg.4775	CN984_27295
Bacillus cereus strain AFS074395	1396.1617	NUZN01000021	fig|1396.1617.peg.1439	COJ77_10215
Bacillus cereus strain AFS096845	1396.1348	NVLK01000072	fig|1396.1348.peg.4895	COM96_25675
Bacillus cereus VD107	1053229.3	AHEX01000016	fig|1053229.3.peg.2133	IIM_02099

Bacillus cereus VD136	1053234.3	AHFC01000058	fig|1053234.3.peg.4199	IIW_02469
Bacillus cereus VDM006	1085379.3	AHFT01000062	fig|1085379.3.peg.4258	KOW_04145
Bacillus cereus VDM021	1085386.3	AHFU01000080	fig|1085386.3.peg.4303	KOY_00793
Bacillus mycoides BHP	1405.8	JMQC01000008	fig|1405.8.peg.299	DJ93_137
Bacillus mycoides Rock1-4	526998.3	NZ_CM000743	fig|526998.3.peg.1633	bmyco0002_29140
Bacillus mycoides strain QHF158	1405.146	RXPJ01000020	fig|1405.146.peg.3094	EKA14_26385
Bacillus mycoides strain SB4	1405.23	MRZX01000031	fig|1405.23.peg.3222	BTH41_03181
Bacillus pseudomycoides strain AFS008599	64104.31	NUDY01000049	fig|64104.31.peg.4277	CN641_19300
Bacillus pseudomycoides strain AFS009893	64104.27	NUDP01000122	fig|64104.27.peg.6021	CN613_26100
Bacillus pseudomycoides strain AFS035137	64104.104	NUTT01000051	fig|64104.104.peg.4025	COF72_12235
Bacillus pseudomycoides strain AFS041167	64104.95	NUQE01000091	fig|64104.95.peg.3852	COE51_20900
Bacillus pseudomycoides strain AFS054612	64104.85	NUQW01000010	fig|64104.85.peg.3479	COE85_03320
Bacillus pseudomycoides strain AFS056801	64104.82	NUQM01000067	fig|64104.82.peg.5196	COE73_21425
Bacillus pseudomycoides strain AFS080374	64104.60	NVGB01000142	fig|64104.60.peg.1123	COL60_26660
Bacillus pseudomycoides strain AFS090198	64104.10	NVOY01000010	fig|64104.10.peg.25	CON64_08925
Bacillus pseudomycoides strain AFS092012	64104.15	NVOR01000007	fig|64104.15.peg.4605	CON65_01490
Bacillus pseudomycoides strain FSL K6-0042	64104.4	MWPX01000040	fig|64104.4.peg.3823	BW425_22870
Bacillus sp. 103mf	1761751.3	FPAF01000006	fig|1761751.3.peg.4294	SAMN04488145_106171
Bacillus sp. 166amftsu	1761753.3	FNQA01000001	fig|1761753.3.peg.1348	SAMN04488156_101145
Bacillus sp. 491mf	1761755.3	FOLV01000020	fig|1761755.3.peg.2310	SAMN04488168_12012
Bacillus sp. AFS015896	2033487.3	NTUO01000045	fig|2033487.3.peg.2809	CN402_15890
Bacillus sp. AFS018417	2033491.3	NTTM01000014	fig|2033491.3.peg.689	CN326_08255
Bacillus sp. AFS019443	2034279.3	NUAA01000022	fig|2034279.3.peg.4807	CN524_05195
Bacillus sp. AFS023182	2033492.3	NTRO01000004	fig|2033492.3.peg.4307	CN288_04745
Bacillus sp. AFS054943	2033506.3	NUGV01000073	fig|2033506.3.peg.4803	CN931_21400

Bacillus sp. AFS098217	2033868.3	NVPR01000016	fig|2033868.3.peg.1444	COO03_04105
Bacillus sp. CDB3 strain CDB3	360310.3	ALBR01000027	fig|360310.3.peg.1974	CDB3_09530
Bacillus sp. FJAT-27238	1679167.3	LGJF01000001	fig|1679167.3.peg.5975	AC624_28195
Bacillus sp. Leaf13	1736211.3	LMRJ01000043	fig|1736211.3.peg.4725	ASG65_26585
Bacillus sp. LF1	1499688.3	CVRB01000002	fig|1499688.3.peg.2616	BN000_02443
Bacillus sp. LK2	1628206.3	LDUK01000085	fig|1628206.3.peg.5766	VK90_27045
Bacillus sp. NH24A2	1866316.3	MAOI01000079	fig|1866316.3.peg.3368	BAU28_00975
Bacillus sp. Soil768D1	1736405.3	LMTA01000055	fig|1736405.3.peg.5253	ASG99_26960
Bacillus thuringiensis serovar navarrensis strain BGSC 4BM1	339658.3	NFDG01000099	fig|339658.3.peg.2653	BK732_12770
Bacillus weihenstephanensis strain SDA_GO95	86662.44	FMAK01000037	fig|86662.44.peg.5367	BWGO95_03270
Bacillus zeae strain JJ-247	1917180.3	QWVT01000021	fig|1917180.3.peg.2058	D1970_12835
bacterium BMS3Abin04	2005712.3	BDSW01000160	fig|2005712.3.peg.1679	BMS3Abin04_01491
Bacteroidales bacterium 43_36	1897038.3	MNQI01000044	fig|1897038.3.peg.3182	BHV67_14415
Bacteroidales bacterium strain UBA12170	2030927.33	DMBV01000110	fig|2030927.33.peg.495	DCG69_07290
Bacteroidales bacterium strain UBA12171	2030927.34	DMBU01000006	fig|2030927.34.peg.697	DCG75_00390
*Bacteroidale*s bacterium strain UBA8399	2030927.40	DMOR01000050	fig|2030927.40.peg.2146	DCQ31_02910
Bacteroides caccae	47678.5	CZBL01000002	fig|47678.5.peg.611	ERS852558_00587
Bacteroides caccae ATCC 43185	411901.7	NZ_AAVM02000004	fig|411901.7.peg.2048	BACCAC_02182
Bacteroides caccae CAG:21	1263037.3	HF997479	fig|1263037.3.peg.3608	BN535_03338
Bacteroides caccae CL03T12C61	997873.3	AGXF01000006	fig|997873.3.peg.1184	HMPREF1061_01109
Bacteroides caccae strain AF24-29LB	47678.165	QRUO01000016	fig|47678.165.peg.1335	DWY26_15800
Bacteroides caccae strain AF46-5GN	47678.173	QRNA01000007	fig|47678.173.peg.4671	DW080_03345
Bacteroides caccae strain AM31-16AC	47678.170	QSJD01000002	fig|47678.170.peg.1301	DW794_01250

Bacteroides faecichinchillae strain DSM 26883	871325.5	FQVD01000034	fig|871325.5.peg.3882	SAMN05444349_13415
Bacteroides faecis MAJ27	1077285.5	FNNN01000006	fig|1077285.5.peg.3335	SAMN05444400_10679
Bacteroides faecis strain OM02-29	674529.16	QSVL01000026	fig|674529.16.peg.2724	DXB21_19540
Bacteroides fragilis 638R	862962.3	FQ312004	fig|862962.3.peg.3138	BF638R_3051
Bacteroides fragilis CAG:47	1263046.3	FR894333	fig|1263046.3.peg.3079	BN669_02834
Bacteroides fragilis CAG:558	1263047.3	HF995768	fig|1263047.3.peg.3026	BN707_02713
Bacteroides fragilis CL05T12C13	997881.3	AGXP01000026	fig|997881.3.peg.2554	HMPREF1080_02433
Bacteroides fragilis CL07T12C05	997883.3	AGXN01000022	fig|997883.3.peg.4015	HMPREF1056_03800
Bacteroides fragilis HMW 615	1073387.4	AGXR01000017	fig|1073387.4.peg.2147	HMPREF1204_02021
Bacteroides fragilis NCTC 9343 strain ATCC 25285	272559.17	NC_003228	fig|272559.17.peg.3250	BF3205
Bacteroides fragilis str. 1007-1-F #3	1339337.3	JGEB01000081	fig|1339337.3.peg.3719	M146_3524
Bacteroides fragilis str. 2-F-2# 4	1339280.3	JGDM01000075	fig|1339280.3.peg.3017	M076_3156
Bacteroides fragilis str. 3725 D9 ii	1339286.3	JNHH01000011	fig|1339286.3.peg.1584	M082_1648
Bacteroides fragilis str. 3774 T13	1339308.3	JGCR01000287	fig|1339308.3.peg.3363	M117_2924
Bacteroides fragilis str. 3783N1-6	1339310.3	JGEU01000040	fig|1339310.3.peg.3462	M119_3184
Bacteroides fragilis str. 3976T8	1339314.3	JGDS01000059	fig|1339314.3.peg.3364	M123_3204
Bacteroides fragilis str. 3986 N(B) 19	1339289.3	JGCW01000299	fig|1339289.3.peg.2487	M085_2713
Bacteroides fragilis str. 3996 N(B) 6	1339283.3	JGDA01000252	fig|1339283.3.peg.3631	M079_3170
Bacteroides fragilis str. 3-F-2 #6	1339335.3	JGDT01000076	fig|1339335.3.peg.2972	M144_3075
Bacteroides fragilis str. DS-208	1339276.3	JGDE01000119	fig|1339276.3.peg.2800	M072_2947
Bacteroides fragilis str. J-143-4	1339271.3	JGDH01000079	fig|1339271.3.peg.2949	M067_3067
Bacteroides fragilis str. Korea 419	1339269.3	JGDW01000075	fig|1339269.3.peg.3792	M065_3882
Bacteroides fragilis str. S23L17	1339329.3	JHEF01000037	fig|1339329.3.peg.2823	M138_2985
Bacteroides fragilis str. S36L11	1339327.3	JGDJ01000246	fig|1339327.3.peg.3656	M136_3103

Bacteroides fragilis str. S6L5	1339318.3	JGVC01000202	fig|1339318.3.peg.4120	M127_3149
Bacteroides fragilis strain 12905	817.252	PDCW01000033	fig|817.252.peg.3908	CQW34_03611
Bacteroides fragilis strain AF14-26	817.340	QRZH01000002	fig|817.340.peg.2182	DWW08_04110
Bacteroides fragilis strain OF05-11AC	817.333	QSWE01000005	fig|817.333.peg.3580	DXA78_10515
Bacteroides fragilis strain Q1F2	817.278	CP018937	fig|817.278.peg.2506	BUN20_11905
Bacteroides fragilis YCH46	295405.11	NC_006347	fig|295405.11.peg.3073	BF3205
Bacteroides nordii CL02T12C05	997884.3	AGXS01000015	fig|997884.3.peg.1826	HMPREF1068_01792
Bacteroides nordii strain AM40-30BH	291645.7	QSGO01000014	fig|291645.7.peg.1003	DW888_15385
Bacteroides ovatus ATCC 8483	411476.11	NZ_DS264560	fig|411476.11.peg.1148	BACOVA_01267
Bacteroides ovatus CAG:22	1263050.3	FR885819	fig|1263050.3.peg.3532	BN541_03267
Bacteroides ovatus CL02T12C04	997885.3	AGXT01000009	fig|997885.3.peg.1742	HMPREF1069_01682
Bacteroides ovatus SD CMC 3f	702443.3	NZ_ADMO01000099	fig|702443.3.peg.3385	CUY_3994
Bacteroides ovatus str. 3725 D1 iv	1339346.3	JNHF01000139	fig|1339346.3.peg.2941	M088_1589
Bacteroides ovatus strain AM17-48	28116.190	QRJR01000008	fig|28116.190.peg.5429	DW206_11695
Bacteroides ovatus strain AM32-14LB	28116.188	QSIX01000022	fig|28116.188.peg.2118	DW803_16385
Bacteroides ovatus strain KLE1656	28116.10	KQ968466	fig|28116.10.peg.4529	HMPREF2532_04520
Bacteroides ovatus strain NLAE-zl-C57	28116.1048	FNDO01000023	fig|28116.1048.peg.3499	SAMN05192582_102349
Bacteroides ovatus strain UBA9483	28116.196	DPUA01000046	fig|28116.196.peg.3797	DHW41_16895
Bacteroides salyersiae CL02T12C01	997887.3	AGXV01000030	fig|997887.3.peg.2486	HMPREF1071_02383
Bacteroides sp. 1_1_14	469585.3	NZ_GG774703	fig|469585.3.peg.815	HMPREF9007_01122
Bacteroides sp. 1_1_30	457387.3	ADCL01000058	fig|457387.3.peg.2907	HMPREF0127_02871
Bacteroides sp. 2_1_22	469588.3	NZ_GG705175	fig|469588.3.peg.2563	HMPREF0102_02455
Bacteroides sp. 2_1_33B	469589.3	NZ_GG705149	fig|469589.3.peg.549	HMPREF0103_0902
Bacteroides sp. 2_2_4	469590.5	NZ_EQ973356	fig|469590.5.peg.1017	BSCG_00874
Bacteroides sp. 3_1_13	457389.4	KQ236956	fig|457389.4.peg.4244	HMPREF9009_04055

Bacteroides sp. 3_1_19	469592.4	NZ_GG774760	fig|469592.4.peg.1756	HMPREF0104_01752
Bacteroides sp. 3_1_23	457390.3	NZ_GG774949	fig|457390.3.peg.1027	HMPREF9010_01842
Bacteroides sp. 41_26	1896973.3	MNQM01000001	fig|1896973.3.peg.374	BHV71_00785
Bacteroides sp. 43_46	1897051.3	MNQN01000014	fig|1897051.3.peg.2494	BHV72_00490
Bacteroides sp. AF16-49	2292192.3	QUHA01000045	fig|2292192.3.peg.3216	DWW69_19380
Bacteroides sp. AF26-7BH	2292193.3	QRTS01000034	fig|2292193.3.peg.3017	DWY71_19545
Bacteroides sp. AF27-33	2292194.3	QRTK01000034	fig|2292194.3.peg.3563	DWY87_22485
Bacteroides sp. AF32-8BH	2302925.3	QVLZ01000013	fig|2302925.3.peg.904	DWZ47_11100
Bacteroides sp. AM16-13	2292938.3	QTLK01000001	fig|2292938.3.peg.210	DW173_01000
Bacteroides sp. AM23-12	2292942.3	QTLF01000002	fig|2292942.3.peg.1829	DW640_03810
Bacteroides sp. AM37-9	2292951.3	QTNF01000005	fig|2292951.3.peg.4206	DW862_05185
Bacteroides sp. AM54-2NS	2292955.3	QTMZ01000021	fig|2292955.3.peg.2158	DXA05_14355
Bacteroides sp. AM56-10ce	2302926.3	QVMA01000058	fig|2302926.3.peg.4852	DXA11_18280
Bacteroides sp. CAG:144	1262736.3	HF999659	fig|1262736.3.peg.57	BN496_00041
Bacteroides sp. CAG:189	1262737.3	HF991029	fig|1262737.3.peg.1149	BN523_01048
Bacteroides sp. CAG:754	1262750.3	FR881775	fig|1262750.3.peg.3362	BN772_03148
Bacteroides sp. CF01-10NS	2292956.3	QTMY01000048	fig|2292956.3.peg.3996	DXA24_12835
Bacteroides sp. D1	556258.5	NZ_EQ973247	fig|556258.5.peg.995	BSAG_01028
Bacteroides sp. D22	585544.3	NZ_GG774822	fig|585544.3.peg.4270	HMPREF0106_04218
Bacteroides sp. HMSC067B03	1739298.3	KV804410	fig|1739298.3.peg.3966	HMPREF2794_18355
Bacteroides sp. HMSC073E02	1739517.3	KV819961	fig|1739517.3.peg.445	HMPREF3015_02095
Bacteroides sp. HPS0048	1078089.3	AGEU01000057	fig|1078089.3.peg.3879	HMPREF1214_03939
Bacteroides sp. KCTC 15687	2447885.3	BHWB01000004	fig|2447885.3.peg.2035	KGMB02408_18340
Bacteroides sp. OF03-11BH	2292957.3	QTMW01000008	fig|2292957.3.peg.5404	DXA54_08180
Bacteroides sp. OM05-10AA	2292282.3	QSUR01000036	fig|2292282.3.peg.3490	DXB58_22125

Bacteroides sp. OM05-12	2292283.3	QSUN01000022	fig|2292283.3.peg.1693	DXB63_14165
Bacteroides thetaiotaomicron	818.28	CZBI01000001	fig|818.28.peg.1061	ERS852557_01027
Bacteroides thetaiotaomicron CAG:40	1263054.3	FR901283	fig|1263054.3.peg.2569	BN644_02357
Bacteroides thetaiotaomicron dnLKV9	1235785.3	ASSM01000008	fig|1235785.3.peg.1801	C799_01804
Bacteroides thetaiotaomicron strain AF37-12	818.297	QROV01000004	fig|818.297.peg.3771	DW011_05150
Bacteroides xylanisolvens CL03T12C04	997892.3	AGXE01000006	fig|997892.3.peg.1081	HMPREF1074_01053
Bacteroides xylanisolvens SD CC 1b	702447.11	CBXG01000049	fig|702447.11.peg.4427	BN890_44450
Bacteroides xylanisolvens strain AF14-7	371601.58	QRYV01000005	fig|371601.58.peg.3545	DWW25_03010
Bacteroides xylanisolvens strain AF38-2	371601.63	QROO01000002	fig|371601.63.peg.1871	DW027_02505
Bacteroides xylanisolvens strain AF39-6AC	371601.62	QROC01000010	fig|371601.62.peg.224	DW042_09170
Bacteroides xylanisolvens strain An107	371601.18	NFLX01000004	fig|371601.18.peg.3457	B5E50_05320
Bacteroides xylanisolvens strain NLAE-zl-C202	371601.392	FOUM01000004	fig|371601.392.peg.2150	SAMN05216250_104178
Bacteroides xylanisolvens strain NLAE-zl-G339	371601.390	FNRP01000009	fig|371601.390.peg.1819	SAMN04487924_109125
Bacteroides xylanisolvens XB1A	657309.4	FP929033	fig|657309.4.peg.3172	BXY_41820
Bacteroides xylanolyticus strain DSM 3808	384636.4	PTJA01000014	fig|384636.4.peg.4422	BXY41_11435
Bacteroidetes bacterium GWF2_41_31	1797348.3	MEOL01000023	fig|1797348.3.peg.1333	A2W85_01605
Bacteroidetes bacterium HGW-Bacteroidetes-16	2013684.3	PHDN01000006	fig|2013684.3.peg.3372	CVT99_04845
Bacteroidetes bacterium strain UBA11986	1898104.116	DPLS01000206	fig|1898104.116.peg.1650	DGH68_05710
Bacteroidetes bacterium strain UBA9647	1898104.106	DOWV01000131	fig|1898104.106.peg.3514	DEP53_06440
Bacteroidetes/Chlorobi group bacterium ChocPot_Mid	2382230.3	RCNQ01000236	fig|2382230.3.peg.2420	D9V86_09360
Blautia coccoides strain NCTC11035	1532.9	UFVT01000001	fig|1532.9.peg.3038	NCTC11035_02878
Blautia hansenii DSM 20583	537007.6	NZ_GG698590	fig|537007.6.peg.2240	BLAHAN_06095
Blautia producta strain DSM 3507	33035.17	PQGB01000420	fig|33035.17.peg.3875	C3R19_23710
Blautia sp. YL58	1796616.4	CP015405	fig|1796616.4.peg.4609	A4V09_20705

Borrelia turicatae 91E135	314724.4	NC_008710	fig|314724.4.peg.168	BT0175
Brachybacterium nesterenkovii strain CIP104813	47847.3	FWFG01000035	fig|47847.3.peg.426	FM110_03950
Brevibacillus agri strain NRRL NRS 1219	51101.6	RHHN01000039	fig|51101.6.peg.2637	EB820_14085
Brevibacillus brevis strain ATCC 35690	1393.7	MXAR01000050	fig|1393.7.peg.5222	B5G50_28710
Brevibacillus brevis strain NCTC2611	1393.21	LR134338	fig|1393.21.peg.1338	NCTC2611_01348
Brevibacillus choshinensis strain DSM 8552	54911.3	LJJB01000007	fig|54911.3.peg.6181	AN963_09815
Brevibacillus fluminis strain JCM 15716	511487.5	RHHQ01000025	fig|511487.5.peg.3190	EDM56_27550
Brevibacillus formosus strain NF2	54913.5	CP018145	fig|54913.5.peg.4088	BP422_18630
Brevibacillus parabrevis strain NRRL NRS 605	54914.7	RHHV01000010	fig|54914.7.peg.492	EDM60_14490
Brevibacillus reuszeri strain DSM 9887	54915.3	LGIQ01000011	fig|54915.3.peg.4749	ADS79_27780
Brevibacillus sp. BC25	1144308.3	AKIX01000130	fig|1144308.3.peg.5905	PMI05_06040
Brevibacillus sp. CF112	1144311.3	AKKB01000101	fig|1144311.3.peg.3427	PMI08_03431
Brevibacillus sp. NRRL NRS-1210	2126352.3	PXZM01000047	fig|2126352.3.peg.5085	C7R93_26020
Brevibacillus sp. NRRL NRS-603	2126351.3	PXZN01000049	fig|2126351.3.peg.4983	C7R94_26505
Brevibacillus sp. strain UBA11289	1882945.5	DOTK01000044	fig|1882945.5.peg.2143	DEP07_12830
Brevibacterium luteolum strain UMB0680	199591.5	PNFZ01000003	fig|199591.5.peg.1617	CJ198_06640
Brevibacterium mcbrellneri ATCC 49030	585530.3	NZ_ADNU01000049	fig|585530.3.peg.1796	HMPREF0183_1839
Brevibacterium paucivorans strain UMB1301	170994.3	PNHK01000003	fig|170994.3.peg.2178	CJ199_09005
Brevibacterium ravenspurgense strain UMB0426	479117.10	PKGO01000008	fig|479117.10.peg.2041	CYJ40_08715
Brevibacterium sp. HMSC063G07	1739261.3	KV802360	fig|1739261.3.peg.1411	HMPREF2757_06965
Brevibacterium sp. HMSC22B09	1581055.3	KV786006	fig|1581055.3.peg.680	HMPREF3087_03435
Brevibacterium sp. HMSC24B04	1581060.5	KV786093	fig|1581060.5.peg.1377	HMPREF3092_06870
Butyribacterium methylotrophicum strain DSM 3468	1487.3	MIMZ01000017	fig|1487.3.peg.2855	BUME_26630

Butyricicoccus pullicaecorum 1.2	1203606.4	AQOB01000012	fig|1203606.4.peg.2564	HMPREF1526_02600
Butyricicoccus pullicaecorum strain An179	501571.10	NFKL01000028	fig|501571.10.peg.1980	B5F15_14990
Butyricicoccus pullicaecorum strain An180	501571.9	NFKK01000023	fig|501571.9.peg.1552	B5F17_13260
Butyricicoccus sp. BB10 strain BB10	1945634.3	NHOC01000005	fig|1945634.3.peg.1833	CBW42_06170
Butyricimonas sp. H184	2093856.3	CP032819	fig|2093856.3.peg.3617	D8S85_17095
Butyricimonas synergistica strain 43_13	544644.3	MNQW01000040	fig|544644.3.peg.729	BHV81_15225
Butyricimonas virosa strain OF02-7	544645.7	QSCR01000025	fig|544645.7.peg.1730	DXA50_13185
Caldilineae bacterium strain J123	2420332.5	RFKH01000616	fig|2420332.5.peg.2819	D6796_09660
Caloranaerobacter azorensis DSM 13643	1121264.4	FQXO01000030	fig|1121264.4.peg.994	SAMN02745135_01318
Caloranaerobacter azorensis H53214	1156417.3	AZTB01000029	fig|1156417.3.peg.1338	Y919_06780
candidate division *Zixibacteria* bacterium strain SURF_9 strain not applicable	2053527.7	QZKQ01000134	fig|2053527.7.peg.2138	C4524_14775
Candidatus Kerfeldbacteria bacterium CG08_land_8_20_14_0_20_42_7	2014245.3	PEXV01000084	fig|2014245.3.peg.301	COT25_02420
Candidatus Komeilibacteria bacterium CG_4_10_14_0_2_um_filter_37_10	1974470.3	PFPO01000021	fig|1974470.3.peg.888	COX77_01125
Candidatus Parcubacteria bacterium strain SURF_31 strain not applicable	2053309.7	QZJU01000006	fig|2053309.7.peg.382	C4546_00855
Candidatus Schekmanbacteria bacterium RBG_13_48_7	1817878.3	MGDD01000264	fig|1817878.3.peg.4021	A2161_21545
Candidatus Schekmanbacteria bacterium RIFCSPLOWO2_12_FULL_38_15	1817883.3	MGDI01000031	fig|1817883.3.peg.2121	A3G31_11505
Candidatus Thorarchaeota archaeon SMTZ-45	1706443.6	LRSM01000148	fig|1706443.6.peg.1870	AM326_03465
Catabacter hongkongensis strain HKU16	270498.26	LAYJ01000111	fig|270498.26.peg.2982	CHK_2030
*Chloroflex*i bacterium RBG_16_50_9	1797643.3	MGNK01000025	fig|1797643.3.peg.455	A2144_01580

Chloroflexi bacterium RBG_19FT_COMBO_47_9	1797665.3	MGOG01000055	fig|1797665.3.peg.673	A2029_07660
Chloroflexi bacterium strain B10_G9	2026724.249	QMOK01000015	fig|2026724.249.peg.474	DRI56_02140
Chloroflexi bacterium strain B27_G4	2026724.246	QMOH01000025	fig|2026724.246.peg.1178	DRI46_05495
Chloroflexi bacterium strain B4_G1	2026724.242	QMNZ01000602	fig|2026724.242.peg.3998	DRI81_17705
Chloroflexi bacterium strain DOLZORAL124_50_6	2026724.107	PDQY01000015	fig|2026724.107.peg.1339	CSB13_01525
Chloroflexi bacterium strain metabat2.725	2026724.297	RPRE01000555	fig|2026724.297.peg.1437	EHM81_10345
Chloroflexi bacterium strain UBA11857	2026724.215	DOTA01000201	fig|2026724.215.peg.2143	DEH25_06780
Chromatiales bacterium strain ARS1279	2026725.3	NZCC01000041	fig|2026725.3.peg.1284	CL797_06150
Clostridiaceae bacterium 14S0207	2082193.3	CP026600	fig|2082193.3.peg.224	C3495_01185
Clostridiales bacterium 38-18	1895746.3	MKTL01000005	fig|1895746.3.peg.1194	BGO41_01065
Clostridiales bacterium 38-18	1895746.3	MKTL01000006	fig|1895746.3.peg.1385	BGO41_12155
Clostridiales bacterium oral taxon 876 str. F0540	1321778.3	AWSZ01000036	fig|1321778.3.peg.1845	HMPREF1982_01871
Clostridiales bacterium oral taxon 876 str. F0540	1321778.3	AWSZ01000058	fig|1321778.3.peg.2933	HMPREF1982_02969
Clostridioides difficile strain 6636-R/ST48	1496.2185	MPEQ01000035	fig|1496.2185.peg.3366	BGU81_16405
Clostridioides difficile strain 7032989	1496.3550	LK933160	fig|1496.3550.peg.3072	BN1095_480085
Clostridioides difficile strain 910500/2012	1496.4103	CAADDZ010000032	fig|1496.4103.peg.3646	SAMEA1710314_03639
Clostridioides difficile strain CD105KSO7	1496.3955	FJUH01000016	fig|1496.3955.peg.999	CDFC105_61004
Clostridioides difficile strain Lei028	1496.3990	CAACZQ010000015	fig|1496.3990.peg.2758	SAMEA1022437_02794
Clostridioides difficile strain lsh12	1496.4001	CAADAA010000023	fig|1496.4001.peg.3584	SAMEA1402358_03637
Clostridium acidurici 9a	1128398.3	CP003326	fig|1128398.3.peg.1996	Curi_c19390
Clostridium acidurici 9a	1128398.3	CP003326	fig|1128398.3.peg.847	Curi_c07960
Clostridium algidicarnis DSM 15099	1121295.3	PTIS01000008	fig|1121295.3.peg.1396	BD821_10826
Clostridium algidixylanolyticum strain SPL73	94868.12	MCIA01000006	fig|94868.12.peg.1088	BET01_14130

Clostridium baratii	1561.29	CZBO01000003	fig|1561.29.peg.1846	ERS852568_01865
Clostridium baratii	1561.29	CZBO01000003	fig|1561.29.peg.2019	ERS852568_02048
Clostridium baratii strain 771-14	1561.27	JZTY01000002	fig|1561.27.peg.1164	UC77_01605
Clostridium baratii strain 771-14	1561.27	JZTY01000010	fig|1561.27.peg.278	UC77_08260
Clostridium bifermentans ATCC 19299	1233170.3	AVNB01000015	fig|1233170.3.peg.1200	C671_1242
Clostridium bifermentans ATCC 638	1233171.3	AVNC01000015	fig|1233171.3.peg.1352	C672_1458
Clostridium botulinum A str. ATCC 3502	413999.7	NC_009495	fig|413999.7.peg.2220	CBO2251
Clostridium botulinum A2 str. Kyoto	536232.3	NC_012563	fig|536232.3.peg.2324	CLM_2458
Clostridium botulinum A3 str. Loch Maree	498214.7	NC_010520	fig|498214.7.peg.2608	CLK_1693
Clostridium botulinum B str. Eklund 17B	508765.6	NC_010674	fig|508765.6.peg.3388	CLL_A3525
Clostridium botulinum B1 str. Okra	498213.7	NC_010516	fig|498213.7.peg.2350	CLD_2324
Clostridium botulinum B2 450	1379739.3	JXSU01000007	fig|1379739.3.peg.2661	N495_11435
Clostridium botulinum Ba4 str. 657	515621.3	NC_012658	fig|515621.3.peg.2665	CLJ_B2462
Clostridium botulinum Bf	445336.4	NZ_ABDP01000014	fig|445336.4.peg.2607	CBB_2494
Clostridium botulinum C str. Eklund	445337.5	NZ_ABDQ01000003	fig|445337.5.peg.951	CBC_A0601
Clostridium botulinum C/D str. BKT12695	1443125.3	JENP01000028	fig|1443125.3.peg.1876	Z962_06120
Clostridium botulinum C/D str. DC5	1443128.3	JDRY01000039	fig|1443128.3.peg.1724	Z955_08805
Clostridium botulinum C/D str. It1	1443126.3	JENO01000049	fig|1443126.3.peg.2061	Z963_08920
Clostridium botulinum D str. 1873	592027.3	NZ_ACSJ01000007	fig|592027.3.peg.1135	CLG_B0976
Clostridium botulinum E1 str. 'BoNT E Beluga'	536233.3	NZ_ACSC01000002	fig|536233.3.peg.17	CLO_0030
Clostridium botulinum F str. Langeland	441772.13	NC_009699	fig|441772.13.peg.2212	CLI_2301
Clostridium botulinum strain 12LNR13-CD	1491.430	LGVT01000008	fig|1491.430.peg.2688	ADU76_00485
Clostridium botulinum strain AM1195	1491.958	CP013701	fig|1491.958.peg.2396	RSJ11_13240
Clostridium botulinum strain AM282	1491.669	CP013683	fig|1491.669.peg.2377	RSJ10_2343
Clostridium botulinum strain ATCC 17786	1491.446	LHUM01000007	fig|1491.446.peg.3626	ADT22_01570

Clostridium botulinum strain ATCC 17862	1491.444	LGIK01000012	fig|1491.444.peg.46	ACP52_01625
Clostridium botulinum strain ATCC 23387	1491.442	LGII01000013	fig|1491.442.peg.94	ACP50_00625
Clostridium botulinum strain CDC_1632	1491.664	CP013243	fig|1491.664.peg.3687	NPD5_3675
Clostridium botulinum strain CDC_67071	1491.733	CP013242	fig|1491.733.peg.776	NPD7_760
Clostridium botulinum strain CDC_67071	1491.736	CP013242	fig|1491.736.peg.776	NPD7_760
Clostridium botulinum strain CDC66008	1491.1060	QVOC01000185	fig|1491.1060.peg.4775	DZC34_06765
Clostridium butyricum 60E.3	997898.4	AGYK01000006	fig|997898.4.peg.2107	HMPREF1084_02088
Clostridium butyricum DORA_1	1403941.3	AZLX01000176	fig|1403941.3.peg.1896	Q607_CBUC00176G0036
Clostridium butyricum E4 str. BoNT E BL5262	632245.3	NZ_ACOM01000005	fig|632245.3.peg.1614	CLP_1039
Clostridium butyricum strain 300064	1492.150	LRDH01000103	fig|1492.150.peg.1926	AWN73_13005
Clostridium butyricum strain SU1	1492.46	LIDW01000003	fig|1492.46.peg.799	AK964_03095
Clostridium carnis strain NCTC10913	1530.3	UYIN01000019	fig|1530.3.peg.1704	NCTC10913_03317
Clostridium cavendishii DSM 21758	1121302.3	FQZB01000013	fig|1121302.3.peg.4037	SAMN02745163_03233
Clostridium celerecrescens 18A	1286362.3	PGET01000001	fig|1286362.3.peg.808	H171_0773
Clostridium chromiireducens strain DSM 23318	225345.3	MZGT01000015	fig|225345.3.peg.1424	CLCHR_14120
Clostridium colicanis 209318	999411.4	AGYT01000007	fig|999411.4.peg.236	HMPREF1092_00253
Clostridium collagenovorans DSM 3089	1121306.4	FQXP01000007	fig|1121306.4.peg.1653	SAMN02745196_02090
Clostridium cylindrosporum DSM 605	1121307.3	LFVU01000024	fig|1121307.3.peg.1862	CLCY_4c02060
Clostridium difficile CD160	1151292.4	AVHW01000040	fig|1151292.4.peg.1476	QEW_1275
Clostridium difficile CD196	645462.3	NC_013315	fig|645462.3.peg.836	CD196_0800
Clostridium difficile NAP08	525259.3	NZ_GG770712	fig|525259.3.peg.1881	HMPREF0220_0665
Clostridium difficile R20291	645463.3	NC_013316	fig|645463.3.peg.840	CDR20291_0781
Clostridium difficile Y384	1151391.4	AVLJ01000432	fig|1151391.4.peg.3930	QQG_1031
Clostridium fallax strain DSM 2631	1533.3	FQVM01000002	fig|1533.3.peg.684	SAMN05443638_10222

Clostridium fallax strain DSM 2631	1533.3	FQVM01000021	fig|1533.3.peg.1123	SAMN05443638_12111
Clostridium felsineum DSM 794 strain DSM 794	1121310.4	LZYT01000005	fig|1121310.4.peg.1450	CLFE_14320
Clostridium frigidicarnis strain DSM 12271	84698.3	FOKI01000005	fig|84698.3.peg.1881	SAMN04488528_1005156
Clostridium gasigenes strain DSM 12272	94869.18	FNJM01000001	fig|94869.18.peg.2532	SAMN04488529_101120
Clostridium grantii DSM 8605	1121316.4	FQXM01000005	fig|1121316.4.peg.2464	SAMN02745207_01232
Clostridium haemolyticum NCTC 8350	1443115.3	JDSA01000126	fig|1443115.3.peg.1103	Z961_05920
Clostridium homopropionicum DSM 5847	1121318.3	LHUR01000012	fig|1121318.3.peg.955	CLHOM_09510
Clostridium hungatei strain DSM 14427	48256.3	MZGX01000009	fig|48256.3.peg.1857	CLHUN_17250
Clostridium intestinale DSM 6191	1121320.3	FQXU01000017	fig|1121320.3.peg.4382	SAMN02745941_04100
Clostridium intestinale URNW	1294142.3	APJA01000012	fig|1294142.3.peg.1189	CINTURNW_1185
Clostridium liquoris strain DSM 100320	1289519.3	PVXO01000047	fig|1289519.3.peg.1797	CLLI_17780
Clostridium neonatale strain LCDC99A005	137838.6	PDCJ01000002	fig|137838.6.peg.3112	CQ394_14905
Clostridium novyi A str. 4552	1444289.3	JENJ01000073	fig|1444289.3.peg.2391	Z968_11750
Clostridium novyi A str. 4570	1444290.3	JDRX01000056	fig|1444290.3.peg.2169	Z969_10815
Clostridium novyi B str. NCTC 9691	1443122.3	JENV01000016	fig|1443122.3.peg.926	Z958_02960
Clostridium novyi NT	386415.7	NC_008593	fig|386415.7.peg.1172	NT01CX_2068
Clostridium novyi strain 150557	1542.10	CP029458	fig|1542.10.peg.1750	DFH04_08985
Clostridium oryzae strain DSM 28571	1450648.3	MZGV01000002	fig|1450648.3.peg.295	CLORY_02830
Clostridium perfringens ATCC 13124	195103.10	NC_008261	fig|195103.10.peg.2808	CPF_2928
Clostridium perfringens ATCC 13124	195103.10	NC_008261	fig|195103.10.peg.291	CPF_0311
Clostridium perfringens B str. ATCC 3626	451754.5	NZ_ABDV01000007	fig|451754.5.peg.1229	AC1_3181
Clostridium perfringens B str. ATCC 3626	451754.5	NZ_ABDV01000014	fig|451754.5.peg.1800	AC1_0337
Clostridium perfringens D str. JGS1721	488537.5	NZ_ABOO01000006	fig|488537.5.peg.672	CJD_0377
Clostridium perfringens D str. JGS1721	488537.5	NZ_ABOO01000039	fig|488537.5.peg.2702	CJD_3250

Clostridium perfringens E str. JGS1987	451755.5	NZ_ABDW01000034	fig|451755.5.peg.3019	AC3_0447
Clostridium perfringens E str. JGS1987	451755.5	NZ_ABDW01000042	fig|451755.5.peg.3284	AC3_3257
Clostridium perfringens SM101	289380.15	NC_008262	fig|289380.15.peg.2532	CPR_2611
Clostridium perfringens SM101	289380.15	NC_008262	fig|289380.15.peg.288	CPR_0306
Clostridium perfringens str. 13	195102.6	NC_003366	fig|195102.6.peg.2672	CPE2601
Clostridium perfringens str. 13	195102.6	NC_003366	fig|195102.6.peg.370	CPE0314
Clostridium perfringens strain AF30-3	1502.480	QRQT01000004	fig|1502.480.peg.2324	DWZ20_05345
Clostridium perfringens strain EHE-NE18	1502.460	CP025501	fig|1502.460.peg.450	CYK96_02415
Clostridium perfringens strain EHE-NE18	1502.460	CP025501	fig|1502.460.peg.860	CYK96_04600
Clostridium perfringens strain JP838	1502.177	CP010994	fig|1502.177.peg.3146	JFP838_15665
Clostridium perfringens strain JP838	1502.177	CP010994	fig|1502.177.peg.334	JFP838_01775
Clostridium perfringens strain MJR7757A	1502.174	KQ956171	fig|1502.174.peg.613	HMPREF3222_00611
Clostridium perfringens strain NCTC10240	1502.490	UFXH01000002	fig|1502.490.peg.2625	NCTC10240_02692
Clostridium perfringens strain NCTC10240	1502.490	UFXH01000002	fig|1502.490.peg.354	NCTC10240_00339
Clostridium perfringens strain PBD1	1502.441	PJST01000003	fig|1502.441.peg.3218	CYK66_11035
Clostridium perfringens strain PBS5	1502.456	PJSS01000002	fig|1502.456.peg.3183	CYK67_11480
Clostridium puniceum strain DSM 2619	29367.3	LZZM01000073	fig|29367.3.peg.1223	CLPUN_11740
Clostridium roseum strain DSM 7320	84029.9	LZYV01000038	fig|84029.9.peg.833	CROST_08240
Clostridium saccharolyticum WM1	610130.3	NC_014376	fig|610130.3.peg.2850	Closa_2625
Clostridium septicum strain DSM 7534	1504.35	CP023671	fig|1504.35.peg.2436	CP523_12225
Clostridium sordellii ATCC 9714	1292036.3	APWR01000093	fig|1292036.3.peg.2229	H477_1499
Clostridium sordellii ATCC 9714	1292036.3	APWR01000093	fig|1292036.3.peg.2875	H477_2156
Clostridium sordellii strain JGS6382	1505.7	LN681234	fig|1505.7.peg.1236	JGS6382_12161
Clostridium sordellii strain JGS6382	1505.7	LN681234	fig|1505.7.peg.840	JGS6382_08261
Clostridium sordellii strain R28058	1505.16	CEKZ01000003	fig|1505.16.peg.1047	R28058_09531

Clostridium sp. 7_2_43FAA	457396.3	NZ_EQ999773	fig|457396.3.peg.2665	CSBG_02653
Clostridium sp. ASBs410	1304866.3	JFBV01000001	fig|1304866.3.peg.4433	K413DRAFT_4415
Clostridium sp. BL8	1354301.3	AUPA01000196	fig|1354301.3.peg.3095	M918_09740
Clostridium sp. CT4	2320868.3	CP025746	fig|2320868.3.peg.2093	C1I91_10210
Clostridium sp. D5	556261.3	NZ_GL870820	fig|556261.3.peg.4366	HMPREF0240_04017
Clostridium sp. HMSC19A10	1581148.4	KV823330	fig|1581148.4.peg.1336	HMPREF3070_06480
Clostridium sp. HMSC19B10	1581182.3	KV785024	fig|1581182.3.peg.1798	HMPREF3074_08765
Clostridium sp. IBUN125C	1523154.3	JZWF01000008	fig|1523154.3.peg.196	ClosIBUN125C_CONTIG16g01020
Clostridium sp. IBUN13A	1523156.3	JZWG01000143	fig|1523156.3.peg.2591	ClosIBUN13A_CONTIG227g03543
Clostridium sp. IBUN22A	1523155.3	JZWE01000131	fig|1523155.3.peg.2862	ClosIBUN22A_CONTIG3g00081
Clostridium sp. JC272	1629550.3	LBBT01000120	fig|1629550.3.peg.523	VN21_05235
Clostridium sp. JN-9	2507159.3	CP035280	fig|2507159.3.peg.879	EQM05_04385
Clostridium sp. K25	1443109.3	JENU01000014	fig|1443109.3.peg.443	Z957_01435
Clostridium sp. Maddingley MBC34-26	1196322.3	ALXI01000057	fig|1196322.3.peg.1594	A370_01633
Clostridium sp. NCR	1538552.3	JQHY01000003	fig|1538552.3.peg.1687	KD33_08460
Clostridium sp. strain UBA9406	1506.54	DNSF01000057	fig|1506.54.peg.292	DDX68_02400
Clostridium sporogenes ATCC 15579	471871.7	NZ_DS981517	fig|471871.7.peg.1314	CLOSPO_02681
Clostridium taeniosporum strain 1/k	394958.4	CP017253	fig|394958.4.peg.2867	BGI42_14260
Clostridium tagluense strain A121	360422.5	BHYK01000004	fig|360422.5.peg.923	Ctaglu_08860
Clostridium tagluense strain A121	360422.5	BHYK01000011	fig|360422.5.peg.2323	Ctaglu_22430
Clostridium tepidiprofundi DSM 19306	1121338.3	LTBA01000004	fig|1121338.3.peg.761	CLTEP_07480
Clostridium termitidis CT1112	1195236.3	AORV01000068	fig|1195236.3.peg.5276	CTER_5081
Clostridium thermobutyricum DSM 4928 strain DSM 4928	1121339.3	LTAY01000048	fig|1121339.3.peg.1903	CLTHE_19160

Clostridium uliginosum strain DSM 12992	119641.3	FOMG01000026	fig|119641.3.peg.729	SAMN05421842_12641
Colwellia marinimaniae strain MTCD1	1513592.3	BDQM01000020	fig|1513592.3.peg.2608	MTCD1_02490
Colwellia sp. MT41	58049.4	CP013145	fig|58049.4.peg.262	CMT41_01280
Coprobacillus sp. 3_3_56FAA	665941.3	ACWL01000072	fig|665941.3.peg.3366	HMPREF1021_03361
Coprobacter sp. strain CIM:MAG 570	1941478.3	QAMD01000012	fig|1941478.3.peg.677	DBY16_04550
Coriobacteriaceae bacterium BV3Ac1	1111135.3	AWUP01000017	fig|1111135.3.peg.1280	HMPREF1248_0104
Coriobacteriia bacterium strain UBA7930	2052159.8	DMLB01000041	fig|2052159.8.peg.342	DCP20_03680
Culturomica sp. strain UBA11491	1926652.3	DOED01000001	fig|1926652.3.peg.879	DD657_00330
Dehalobacter sp. TeCB1	1843715.4	MCHF01000016	fig|1843715.4.peg.1014	A7D23_08360
Dehalococcoidia bacterium SG8_51_3	1703394.4	LJTX01000017	fig|1703394.4.peg.1504	AMJ70_01620
Deinococcus sp. K2S05-167	2320857.3	QYUJ01000014	fig|2320857.3.peg.384	D3875_05845
Dermatophilus congolensis strain NCTC13039	1863.3	LT906453	fig|1863.3.peg.1077	SAMEA4475696_01042
Dermatophilus congolensis strain NCTC7915	1863.4	UFYA01000001	fig|1863.4.peg.442	NCTC7915_00413
Desulfosporosinus sp. Tol-M	1536651.3	JQID01000203	fig|1536651.3.peg.2958	JT05_14330
Desulfotomaculum guttoideum strain DSM 4024	58134.4	FOIP01000001	fig|58134.4.peg.5736	SAMN05443270_3378
Desulfuromonadales bacterium strain GT-UBC1	2099678.6	RHLS01000006	fig|2099678.6.peg.892	ED859_04680
Dethiosulfatibacter aminovorans DSM 17477	1121476.3	FQZL01000022	fig|1121476.3.peg.1459	SAMN02745751_02688
Dorea sp. 5-2	1235798.3	ASTD01000024	fig|1235798.3.peg.1295	C817_01216
Emergencia timonensis strain AM07-24	1776384.4	QRMS01000002	fig|1776384.4.peg.1521	DW099_06585
Enterococcus asini ATCC 700915 [PRJNA202676]	1158606.3	AJAP01000019	fig|1158606.3.peg.1861	UAS_01917
Enterococcus faecium strain Isolate 3	1352.3639	FKLT01000056	fig|1352.3639.peg.4953	DTPHA_1402985
Enterococcus pallens ATCC BAA-351 [PRJNA202677]	1158607.3	AJAQ01000001	fig|1158607.3.peg.637	UAU_00633
Enterococcus phoeniculicola ATCC BAA-412 [PRJNA202680]	1158610.3	AJAT01000018	fig|1158610.3.peg.3302	UC3_03311

Enterococcus raffinosus ATCC 49464 [PRJNA202673]	1158602.3	AJAL01000012	fig|1158602.3.peg.2482	UAK_02481
Enterococcus rivorum strain LMG 25899	762845.3	MIEK01000078	fig|762845.3.peg.3560	BCR26_06435
Enterococcus thailandicus strain a523	417368.9	CP023074	fig|417368.9.peg.1699	CK496_08255
Enterococcus thailandicus strain F0711D 46	417368.5	LWMN01000001	fig|417368.5.peg.227	A6E74_01120
Erysipelatoclostridium ramosum strain AM21-17	1547.15	QUSJ01000014	fig|1547.15.peg.970	DW242_12010
Eubacterium aggregans strain SR12	81409.3	FNRK01000023	fig|81409.3.peg.1880	SAMN04515656_12317
Eubacterium angustum strain DSM 1989	39480.3	MKIE01000009	fig|39480.3.peg.1971	EUAN_20010
Eubacterium barkeri strain VPI 5359	1528.7	FNOU01000005	fig|1528.7.peg.1786	SAMN04488579_10568
Eubacterium callanderi strain NLAE-zl-G225	53442.3	FOWI01000006	fig|53442.3.peg.1609	SAMN04487888_106131
Eubacterium combesii strain DSM 20696	39481.3	PEIK01000007	fig|39481.3.peg.3117	CS538_10030
Eubacterium contortum	39482.3	CYZU01000057	fig|39482.3.peg.4633	ERS852491_04276
Eubacterium contortum strain 2789STDY5834876	39482.4	CYZU01000057	fig|39482.4.peg.4634	ERS852491_04276
Eubacterium limosum KIST612	903814.3	NC_014624	fig|903814.3.peg.1062	ELI_1157
Eubacterium limosum strain 8486cho	1736.15	QGUD01000011	fig|1736.15.peg.3640	C7955_11192
Eubacterium limosum strain SA11	1736.8	CP011914	fig|1736.8.peg.900	ACH52_0856
Eubacterium sp. AM05-23	2292043.3	QUDF01000016	fig|2292043.3.peg.1361	DW091_18115
Eubacterium sp. YI	2041044.3	NXNL01000005	fig|2041044.3.peg.155	CPZ25_15875
Firmicutes bacterium HGW-Firmicutes-15	2013776.3	PHAA01000008	fig|2013776.3.peg.962	CVU90_05060
Firmicutes bacterium HGW-Firmicutes-4	2013785.3	PGZR01000006	fig|2013785.3.peg.2310	CVU99_01900
Flavobacteriaceae bacterium strain UBA7949	1871037.131	DNYW01000209	fig|1871037.131.peg.1142	DDZ39_06835
Flavobacteriaceae bacterium strain UBA7949	1871037.131	DNYW01000291	fig|1871037.131.peg.1379	DDZ39_09820
Fusobacteria bacterium strain MAG 22	2060921.4	QNYN01000075	fig|2060921.4.peg.628	DSY38_01970
Gammaproteobacteria bacterium strain B25_G4	1913989.499	QNFD01000371	fig|1913989.499.peg.2937	DRQ59_13725
Gammaproteobacteria bacterium strain NORP60	1913989.13	NVVP01000023	fig|1913989.13.peg.1000	COA90_06085

Gammaproteobacteria bacterium strain NORP60	1913989.13	NVVP01000023	fig|1913989.13.peg.1011	COA90_06130
Gammaproteobacteria bacterium strain UBA7956	1913989.398	DOIT01000330	fig|1913989.398.peg.1359	DD827_09815
Gemmatimonas sp. SG8_17	1703355.3	LJNM01000328	fig|1703355.3.peg.2047	AMS18_14140
Gemmatimonas sp. SG8_28	1703357.3	LJNQ01000031	fig|1703357.3.peg.2515	AMS20_02830
Gemmiger sp. An194 strain An194	1965582.3	NFKA01000003	fig|1965582.3.peg.1962	B5F28_03140
Gemmiger sp. An50 strain An50	1965639.3	NFID01000005	fig|1965639.3.peg.3038	B5G03_06750
Gemmiger sp. An87 strain An87	1965662.3	NFHD01000002	fig|1965662.3.peg.1211	B5G38_01815
Geobacter bemidjiensis Bem	404380.4	NC_011146	fig|404380.4.peg.3391	Gbem_3468
Geobacter daltonii FRC-32	316067.3	NC_011979	fig|316067.3.peg.1912	Geob_1856
Geobacter metallireducens GS-15	269799.8	NC_007517	fig|269799.8.peg.2418	Gmet_2380
Geobacter pelophilus strain Drf2	60036.3	BDQG01000001	fig|60036.3.peg.3663	GPEL0_01f5365
Geobacter sp. DSM 9736 strain DSM 9736	1277350.3	LT896716	fig|1277350.3.peg.394	SAMN06269301_0377
Geobacter sp. GSS01	1510391.3	JXBL01000001	fig|1510391.3.peg.2215	SE37_09025
Geobacter sp. M18	443143.4	NC_014973	fig|443143.4.peg.597	GM18_0606
Geobacter sp. M21	443144.3	NC_012918	fig|443144.3.peg.3464	GM21_3545
Geobacter sp. OR-1	1266765.3	BAZF01000001	fig|1266765.3.peg.295	OR1_00280
Geobacter sp. strain H2geo	46610.14	PJFC01000010	fig|46610.14.peg.1013	CXR31_15280
Geobacter sp. strain L1geo	46610.15	PJFB01000001	fig|46610.15.peg.142	CXR30_00665
Geobacter sp. strain UBA9964	46610.11	DNJC01000115	fig|46610.11.peg.1358	DCZ75_15295
Geobacter sp. strain UBA9964	46610.11	DNJC01000115	fig|46610.11.peg.1364	DCZ75_15325
Geobacter sp. strain UBA9976	46610.12	DNRQ01000164	fig|46610.12.peg.2402	DDY22_16205
Geobacter sp. strain UBA9976	46610.12	DNRQ01000164	fig|46610.12.peg.2417	DDY22_16265
Geobacter sulfurreducens PCA	243231.5	NC_002939	fig|243231.5.peg.1191	GSU1196
Geobacter sulfurreducens strain AM-1	35554.4	CP010430	fig|35554.4.peg.3217	RW64_15680
Geobacter sulfurreducens strain YM18	35554.10	AP017912	fig|35554.10.peg.1188	YM18_1177

Geobacter uraniireducens Rf4	351605.6	NC_009483	fig|351605.6.peg.3516	Gura_3294
Geobacteraceae bacterium GWC2_48_7	1798315.3	MGZI01000053	fig|1798315.3.peg.638	A2079_06935
Geobacteraceae bacterium GWC2_58_44	1798318.3	MGZL01000098	fig|1798318.3.peg.3680	A2075_22065
Geobacteraceae bacterium GWC2_58_44	1798318.3	MGZL01000098	fig|1798318.3.peg.3694	A2075_22125
Hungatella hathewayi strain AF19-21	154046.44	QVIA01000003	fig|154046.44.peg.2950	DWX41_03975
Ignavibacteria bacterium CG2_30_36_16	1805221.3	MNYQ01000041	fig|1805221.3.peg.3274	AUK34_02635
Ignavibacteria bacterium RBG_13_36_8	1798432.3	MHAE01000069	fig|1798432.3.peg.1167	A2V66_01435
Ignavibacteria bacterium strain BM516	2053306.3	PKTF01000016	fig|2053306.3.peg.1059	C0600_00500
Ignavibacteriales bacterium strain UBA8501	2049428.7	DPRR01000014	fig|2049428.7.peg.1851	DHV28_06210
Klebsiella pneumoniae CHS 70	1438774.3	JMYN01000002	fig|1438774.3.peg.718	AF26_00709
Ktedonobacterales bacterium Uno11	2014872.3	BIFS01000002	fig|2014872.3.peg.8695	KDK_76910
Kurthia sp. 11kri321	1750719.3	CP013217	fig|1750719.3.peg.2734	ASO14_2774
Lachnoclostridium sp. strain UBA11633	2028282.7	DPAR01000371	fig|2028282.7.peg.4009	DEQ64_19060
Lachnoclostridium sp. strain UBA11745	2028282.4	DNNH01000149	fig|2028282.4.peg.2889	DC053_12975
Lachnospiraceae bacterium 6_1_63FAA	658083.3	NZ_ACTV01000015	fig|658083.3.peg.1333	HMPREF0992_01315
Lachnospiraceae bacterium AM25-39	2302972.3	QVHF01000014	fig|2302972.3.peg.749	DW684_11155
Lachnospiraceae bacterium CAG:364 strain MGS:364	1262983.4	FR888143	fig|1262983.4.peg.1338	BN627_01235
Lachnospiraceae bacterium strain UBA9840	1898203.25	DMLD01000007	fig|1898203.25.peg.903	DCP06_00340
Lactobacillus buchneri CD034	1071400.3	CP003043	fig|1071400.3.peg.424	LBUCD034_0431
Lactobacillus diolivorans DSM 14421	1423739.3	AZEY01000098	fig|1423739.3.peg.1395	FC85_GL001332
Lactobacillus kefiri DSM 20587 = JCM 5818	1423764.5	AYYV01000031	fig|1423764.5.peg.1180	FC95_GL001161
Lactobacillus kefiri strain OG2	33962.5	NCWS01000009	fig|33962.5.peg.2687	B8W85_03910
Lactobacillus otakiensis JCM 15040	1291737.4	BASH01000002	fig|1291737.4.peg.774	LOT_0779
Lactobacillus parafarraginis DSM 18390 = JCM 14109	1423786.4	AZFZ01000008	fig|1423786.4.peg.2901	FD47_GL002757

Lactobacillus parafarraginis F0439	797515.3	AGEY01000105	fig|797515.3.peg.1716	HMPREF9103_01861
Lactobacillus parakefiri DSM 10551	1423787.3	AZEN01000052	fig|1423787.3.peg.3751	FD08_GL003520
Lactobacillus rapi DSM 19907 = JCM 15042	1423795.5	AZEI01000076	fig|1423795.5.peg.683	FD12_GL000671
Lactobacillus sp. strain LAC4	1591.25	QOUP01000002	fig|1591.25.peg.526	DUD34_02555
Lactobacillus sunkii DSM 19904	1423808.3	AZEA01000010	fig|1423808.3.peg.438	FD17_GL000432
Leifsonia pindariensis strain PON 10	478010.4	MPZN01000041	fig|478010.4.peg.2454	GY24_12030
Leifsonia sp. strain NAT116	1870902.3	PABQ01000044	fig|1870902.3.peg.2213	CMF56_11870
Leifsonia sp. strain NP56	1870902.4	PBTR01000083	fig|1870902.4.peg.1915	CMF57_09425
Longilinea arvoryzae strain KOME-1	360412.5	DF967973	fig|360412.5.peg.4081	LARV_03911
Lysinibacillus sp. BF-4	1473546.3	JPUW01000013	fig|1473546.3.peg.1240	CH76_06305
Lysinibacillus sp. FJAT-14745	1704289.4	LITM01000008	fig|1704289.4.peg.4779	AMS59_15245
Lysinibacillus sp. GY32	1145276.3	CP006837	fig|1145276.3.peg.3289	T479_16725
Lysinibacillus sphaericus strain A1	1421.61	PGLV01000001	fig|1421.61.peg.251	LYSIN_00247
Lysinibacillus sphaericus strain DSM 28	1421.62	CP019980	fig|1421.62.peg.1376	LS41612_06975
Maribacter polysiphoniae strain DSM 23514	429344.6	QGGQ01000003	fig|429344.6.peg.2074	LX92_01946
Maribacter sp. strain DOLZORAL124_41_26	1897614.3	PDPR01000014	fig|1897614.3.peg.1626	CR994_06820
Marinilabiliales bacterium strain BM718	2053303.3	PKTC01000062	fig|2053303.3.peg.1575	C0597_02245
Melissococcus sp. OM08-11BH	2293110.3	QTZC01000006	fig|2293110.3.peg.1841	DXC12_08680
Methylophaga sp. strain NORP53	2024840.12	NVVW01000012	fig|2024840.12.peg.2302	COA83_06495
Methylophaga sp. strain NORP65	2024840.11	NVVK01000002	fig|2024840.11.peg.105	COA95_01670
Methylophaga sp. strain NORP65	2024840.11	NVVK01000002	fig|2024840.11.peg.106	COA95_01675
Microbacterium aurum strain KACC 15219	36805.5	CP018762	fig|36805.5.peg.1269	BOH66_05925
Microbacterium hominis strain SJTG1	162426.5	CP025299	fig|162426.5.peg.2438	CXR34_11990
Microbacterium sp. 67-17	1895782.3	MKTR01000014	fig|1895782.3.peg.1185	BGO47_06985
Microbacterium sp. AG790	2183995.3	RBWZ01000008	fig|2183995.3.peg.1488	DEU37_1619

Microbacterium sp. PM5	2014534.3	CP022162	fig|2014534.3.peg.1335	CEP17_06675
Microbacterium sp. strain EAC103	51671.11	NZTV01000112	fig|51671.11.peg.479	CMH34_13625
Microcella sp. HL-107	2035245.3	PHUH01000001	fig|2035245.3.peg.309	CLT70_0310
Microlunatus phosphovorus NM-1	1032480.4	NC_015635	fig|1032480.4.peg.4381	MLP_44290
Micropruina glycogenica strain 1	75385.5	LT985188	fig|75385.5.peg.3326	MPLG2_3418
Microterricola viridarii strain DSM 21772	412690.4	LT629742	fig|412690.4.peg.2085	SAMN04489834_2083
Microterricola viridarii strain ERGS5:02	412690.6	CP014145	fig|412690.6.peg.1275	AWU67_06240
Mobilicoccus pelagius NBRC 104925	1089455.3	BAFE01000001	fig|1089455.3.peg.63	MOPEL_001_00640
Mycobacterium europaeum strain CSUR P1344	761804.3	CTEC01000001	fig|761804.3.peg.2495	BN000_02443
Oceanithermus profundus DSM 14977	670487.3	NC_014761	fig|670487.3.peg.213	Ocepr_0095
Odoribacter sp. CAG:788	1262909.3	HF993872	fig|1262909.3.peg.778	BN783_00704
Odoribacter sp. OF09-27XD	2293112.3	QUMF01000048	fig|2293112.3.peg.2922	DXA95_13925
Odoribacter splanchnicus CAG:14	1263090.3	FR882811	fig|1263090.3.peg.882	BN493_00808
Odoribacter splanchnicus DSM 220712	709991.3	NC_015160	fig|709991.3.peg.1802	Odosp_1738
Odoribacter splanchnicus strain AF14-6AC	28118.35	QRYW01000034	fig|28118.35.peg.2510	DWW24_14775
Paenibacillus sp. 11 strain 11	1852522.3	FXAZ01000002	fig|1852522.3.peg.2360	SAMN06295960_2175
Paeniclostridium sordellii 8483	1172204.5	AJXR02000029	fig|1172204.5.peg.1421	WS9_008635
Paeniclostridium sordellii strain UMC2	1505.67	CDLK01000002	fig|1505.67.peg.1533	UMC2_13681
Parabacteroides distasonis ATCC 8503	435591.13	NC_009615	fig|435591.13.peg.968	BDI_0985
Parabacteroides distasonis str. 3776 Po2 i	1339341.3	JNHL01000113	fig|1339341.3.peg.2707	M090_4100
Parabacteroides distasonis str. 3999B T(B) 4	1339344.3	JNHP01000158	fig|1339344.3.peg.3554	M095_3836
Parabacteroides distasonis strain AF36-3	823.236	QRPA01000001	fig|823.236.peg.131	DW002_00640
Parabacteroides distasonis strain AM16-4	823.233	QRKE01000008	fig|823.233.peg.4624	DW188_11940
Parabacteroides distasonis strain An199	823.84	NFJX01000008	fig|823.84.peg.4285	B5F32_10945
Parabacteroides distasonis strain CBA7138	823.214	NNCA01000003	fig|823.214.peg.4244	CF162_20325

Parabacteroides distasonis strain OF04-11BH	823.221	QSCD01000017	fig|823.221.peg.1267	DXA71_10320
Parabacteroides distasonis strain UBA11942	823.239	DOCE01000081	fig|823.239.peg.1569	DD441_05445
Parabacteroides gordonii MS-1	1203610.3	KQ033920	fig|1203610.3.peg.4566	HMPREF1536_04479
Parabacteroides sp. 20_3 strain TF09-4	469591.9	QSQY01000011	fig|469591.9.peg.783	DXC95_12800
Parabacteroides sp. AF19-14	2293114.3	QTMJ01000013	fig|2293114.3.peg.865	DWX33_13650
Parabacteroides sp. AF27-14	2293116.3	QTLZ01000017	fig|2293116.3.peg.1264	DWY79_15905
Parabacteroides sp. AF39-10AC	2293117.3	QTLP01000001	fig|2293117.3.peg.15	DW033_00065
Parabacteroides sp. AM44-16	2293122.3	QTNC01000005	fig|2293122.3.peg.4206	DW945_06705
Parabacteroides sp. CH2-D42-20	2320086.3	RAYG01000005	fig|2320086.3.peg.3374	D7V92_04775
Parabacteroides sp. CT06	2025876.3	CP022754	fig|2025876.3.peg.1777	CI960_08640
Parabacteroides sp. D13	563193.3	NZ_GG698739	fig|563193.3.peg.1635	HMPREF0619_01632
Parabacteroides sp. HGS0025	1078087.3	KQ033902	fig|1078087.3.peg.1966	HMPREF1212_01946
Parabacteroides sp. OF01-14	2293123.3	QTMX01000001	fig|2293123.3.peg.370	DXA29_01835
Parabacteroides sp. OF04-13BH	2293124.3	QTMV01000008	fig|2293124.3.peg.4529	DXA72_08805
Paraclostridium bifermentans strain SU1074NT	1490.6	MWJI01000001	fig|1490.6.peg.403	B2H97_01970
Pelolinea submarina strain DSM 23923	913107.6	QUMS01000006	fig|913107.6.peg.3212	DFR64_3140
Pelosinus fermentans B4	1149862.3	AKVJ01000076	fig|1149862.3.peg.4229	FB4_1206
Pelosinus propionicus DSM 13327	1123291.3	FOTS01000007	fig|1123291.3.peg.1419	SAMN04490355_100791
Peptoclostridium difficile ATCC 9689 = DSM 1296	1121308.8	CP011968	fig|1121308.8.peg.918	CDIF1296T_01018
Peptoclostridium difficile strain RA09_70	1496.848	JPPA01000036	fig|1496.848.peg.941	IM33_04070
Peptostreptococcus russellii strain Calf135	215200.11	FODF01000002	fig|215200.11.peg.953	SAMN05216454_10259
Porphyromonadaceae bacterium strain UBA11471	2049046.39	DPQM01000025	fig|2049046.39.peg.2270	DHU85_06745
Porphyromonadaceae bacterium strain UBA12018	2049046.4	DLYT01000012	fig|2049046.4.peg.2216	DCF91_01495

Porphyromonas sp. 31_2	658663.3	ACUD01000002	fig|658663.3.peg.1143	HMPREF1002_01095
Propionibacteriaceae bacterium NML 130396	2016507.3	NMVQ01000001	fig|2016507.3.peg.482	CGZ93_02330
Propionibacteriaceae bacterium strain UBA11038	2021380.10	DORN01000071	fig|2021380.10.peg.2536	DEG88_04695
Propionibacteriaceae bacterium strain UBA11038	2021380.10	DORN01000157	fig|2021380.10.peg.3599	DEG88_10435
Propionibacteriaceae bacterium strain UBA8946	2021380.8	DMIX01000093	fig|2021380.8.peg.1269	DCM67_02850
Propionicimonas paludicola strain DSM 15597	185243.3	PDJC01000001	fig|185243.3.peg.1002	ATK74_1001
Proteobacteria bacterium strain DOLZORAL124_48_12	1977087.21	PDPG01000068	fig|1977087.21.peg.1805	CSB47_11335
Pseudoalteromonas sp. NBT06-2	2025950.3	NQMR01000004	fig|2025950.3.peg.3621	CJF42_01080
Pseudonocardia ammonioxydans strain CGMCC 4.1877	260086.4	FOUY01000034	fig|260086.4.peg.1398	SAMN05216207_103423
Rhodococcus agglutinans strain CCTCC AB2014297	1644129.3	RKLP01000005	fig|1644129.3.peg.3667	EGT67_11075
Rhodococcus sp. 1163 strain 1163	1905289.3	MKKX01000033	fig|1905289.3.peg.3014	BJD99_11470
Rhodococcus sp. AG1013	2183996.3	QQAT01000019	fig|2183996.3.peg.4425	DEU38_119115
Rhodococcus sp. AQ5-07	2054902.3	KZ845637	fig|2054902.3.peg.5579	CVN56_26850
Rhodococcus sp. OK270	1882814.3	OBDL01000008	fig|1882814.3.peg.1328	SAMN05447004_10891
Rhodococcus sp. OK302	1882769.3	NPJZ01000001	fig|1882769.3.peg.4179	BDB13_4069
Rhodococcus sp. RD6.2	260936.4	CVQP01000006	fig|260936.4.peg.2365	RHCRD62_30246
Rhodococcus tukisamuensis strain JCM 11308	168276.3	FNAB01000003	fig|168276.3.peg.3195	SAMN05444580_103361
Rikenella microfusus strain NCTC11190	28139.3	UGVL01000001	fig|28139.3.peg.1920	NCTC11190_01633
Robiginitomaculum sp. strain NORP13	2030823.12	NVXK01000009	fig|2030823.12.peg.1772	COA43_07640
Robiginitomaculum sp. strain NORP161	2030823.4	NVRS01000002	fig|2030823.4.peg.650	COB92_00945
Robiginitomaculum sp. strain NORP172	2030823.3	NVRH01000012	fig|2030823.3.peg.550	COC03_05195
Robiginitomaculum sp. strain NORP39	2030823.9	NVWK01000004	fig|2030823.9.peg.2723	COA69_06610

Robinsoniella sp. RHS	1504536.19	JNGB01000024	fig|1504536.19.peg.3398	RHS_2825
Romboutsia lituseburensis DSM 797	1121325.3	FNGW01000010	fig|1121325.3.peg.1465	SAMN04515677_1107
Romboutsia sp. Frifi strain FRIFI	1507512.3	LN650648	fig|1507512.3.peg.354	FRIFI_0356
Salinibacterium amurskyense strain DSM 16400	205941.3	PGFH01000001	fig|205941.3.peg.4	CLV85_0004
Salinibacterium sp. UTAS2018	2508880.3	CP035375	fig|2508880.3.peg.2246	ESZ53_11085
Spirochaeta sp. strain UBA9216	28185.4	DMKD01000256	fig|28185.4.peg.1617	DCO79_09630
Spirochaetae bacterium HGW-Spirochaetae-7	2013839.3	PGXR01000196	fig|2013839.3.peg.2854	CVV51_12380
Spirochaetes bacterium GWB1_59_5	1802176.3	MIAQ01000113	fig|1802176.3.peg.1454	A2Y38_09095
Sporomusa acidovorans DSM 3132	1123286.3	LSLL01000055	fig|1123286.3.peg.2860	SPACI_26670
Sporomusa malonica strain DSM 5090	112901.3	FWXI01000004	fig|112901.3.peg.4499	SAMN04488500_10474
Sporomusa silvacetica DSM 10669	1123289.3	LSLK01000057	fig|1123289.3.peg.1263	SPSIL_11490
Sporomusa sp. An4 strain *Sporomusa ovata* strain An4	411922.4	CTRP01000012	fig|411922.4.peg.3314	SpAn4DRAFT_2385
Sporosarcina sp. P16a	2048262.4	PDYN01000003	fig|2048262.4.peg.1682	CSV78_04365
Sporosarcina sp. P16a	2048262.4	PDYN01000003	fig|2048262.4.peg.1644	CSV78_04185
Staphylococcus massiliensis S46	1229783.3	AMSQ01000003	fig|1229783.3.peg.543	C273_02678
Syntrophomonas sp. strain UBA11028	2053627.5	DOIG01000070	fig|2053627.5.peg.276	DD791_07695
Terrisporobacter glycolicus strain KPPR-9	36841.13	FORW01000001	fig|36841.13.peg.2491	SAMN02910355_2517
Terrisporobacter glycolicus strain UBA8115	36841.11	DNVD01000092	fig|36841.11.peg.816	DDY58_06875
Terrisporobacter sp. 08-306576	1577792.3	JWHR01000064	fig|1577792.3.peg.1223	QX51_05965
Tessaracoccus sp. OH4464_COT-324	2491059.3	RQYZ01000004	fig|2491059.3.peg.1793	EII42_03605
uncultured *Clostridium* sp. strain 2789STDY5834859	59620.19	FMFC01000006	fig|59620.19.peg.1594	SAMEA3545325_02527
uncultured *Clostridium* sp. strain 2789STDY5834951	59620.36	FMGR01000001	fig|59620.36.peg.2368	SAMEA3545414_01183
uncultured *Clostridium* sp. strain 2789STDY5834951	59620.36	FMGR01000001	fig|59620.36.peg.2144	SAMEA3545414_00955
uncultured *Clostridium* sp. strain 2789STDY5834953	59620.38	FMGT01000038	fig|59620.38.peg.256	SAMEA3545416_06137

uncultured *Eubacterium* sp. strain 2789STDY5834905	165185.11	FMHH01000007	fig|165185.11.peg.2814	SAMEA3545368_03453
Vagococcus fessus strain CCUG 41755	120370.3	NGJY01000001	fig|120370.3.peg.803	CBF31_04020
Vagococcus humatus strain JCM 31581	1889241.3	PXZH01000004	fig|1889241.3.peg.1564	C7P63_07960
Vagococcus lutrae LBD1	1408226.3	AYSH01000019	fig|1408226.3.peg.1489	T233_01534
Vagococcus lutrae strain CCUG 39187	81947.4	NGJW01000001	fig|81947.4.peg.269	CBF33_01340
Vagococcus salmoninarum strain NCFB 2777	2739.3	NGJU01000030	fig|2739.3.peg.2049	CBF35_14240
Vagococcus sp. D7T301 strain D7T301	1768210.4	MVAB01000001	fig|1768210.4.peg.1245	BW731_06255
Vagococcus sp. SS1994	1977868.3	NGJT01000013	fig|1977868.3.peg.500	CBF36_07830