The Chemical Biology of Siderophores from the Bacterial Endosymbionts of Shipworms

By

Andrew W. Han

A dissertation presented to the
Division of Environmental & Biomolecular Systems
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Environmental Science and Engineering

November 2011

DIVISION OF ENVIRONMENTAL AND BIOMOLECULAR SYSTEMS
INSTITUTE OF ENVIRONMENTAL HEALTH
OREGON HEALTH & SCIENCE UNIVERSITY

CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. dissertation of

Andrew W. Han

has been approved

Dr. Peter Zuber, Professor
Examination Committee Chair

Dr. Margo G. Haygood, Professor
Research Advisor

Dr. Jorge Crosa, Professor
Dept. of Molecular Microbiology and Immunology
External Examiner

Dr. Alison Butler, Professor
Dept. of Chemistry and Biochemistry
University of California, Santa Barbara
External Examiner

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to first and foremost thank my advisor, Dr. Margo Haygood, for the opportunity to study in her lab. During my time with her, she has granted me the freedom to explore my various interests and has helped guide me along each step of the way.

I would also like to thank Dr. Brad Tebo and Carolyn Sheehan who, along with Dr. Haygood, have created the super-group known as the Haybo Lab and have helped to facilitate the open dialogue amongst the group. I want to extend a special thank you directly to Carolyn Sheehan for facilitating all of my needs during my graduate studies and, of course, being my dive buddy. To all the past and present members of the Haygood-Tebo lab, I appreciate the helpful discussions, troubleshooting, and coffee breaks that have helped get to this point.

I would like to thank Dr. Holly Simon and her lab for the support and insight they have provided through our weekly group meetings. I would like to thank Dr. Alison Butler and Moriah Sandy for their collaboration and assistance in the structural elucidation of siderophores. I would like to thank Dr. Jorge Crosa and Dr. Hiroaki Naka for their help in developing a mutagenesis protocol in *Teredinibacter turnerae*. They provided me with the necessary strains, plasmids, and support to get me started.

I would like to thank Nancy Christie who has generally made sure that I'm not doing anything wrong, academically and administratively.

I want to thank all my committee members for taking the time to read and evaluate my graduate work.

And I want to thank all my family and friends for their support through this journey. I want to thank my grandparents for the constant and unending words of encouragement and prayers. And most importantly, I want to thank my parents, Elizabeth, Charline, and Kelly, all of whom without I would not be at this point.

This work was funded in part by NIH grant U01TW008163.

Abstract

The Chemical Biology of Siderophores from the Bacterial Endosymbionts of Shipworms

By
Andrew Han
Doctor of Philosophy
Oregon Health & Science University
November 2011
Professor Margo Haygood, Advisor

Shipworms are marine bivalve mollusks (Family Teredinidae) that burrow into wood for shelter and a source of food. They harbor a closely related, yet phylogenetically distinct, group of bacterial endosymbionts in bacteriocytes located in the gills. This endosymbiotic community is believed to support the host's nutrition in multiple ways. The symbionts have been shown to produce cellulolytic enzymes that are believed to aid the host in the digestion of wood. They also provide a source of fixed nitrogen through the process of nitrogen fixation to supplement the nitrogen-poor diet of wood. In host systems, iron is tightly regulated and bound by host proteins, limiting its availability to associated microbes. Due to their implicated role in nitrogen fixation, shipworm symbionts may have an even higher demand for iron in the host beyond those required for normal growth. One of the strategies utilized by microbes in the acquisition of iron is through the production of low molecular weight compounds with high affinity for Fe(III) called siderophores.

This work presents the structures, biosynthesis, and possible roles for siderophores in the shipworm system. One of the most well studied shipworm symbiont is *Teredinibacter turnerae*. Chapter 2 presents the characterization of a novel catecholate siderophore produced by *T. turnerae* T7901, called turnerbactin. In addition, the biosynthetic pathway for turnerbactin is analyzed. Chapter 3 introduces

other members of the shipworm endosymbiotic community and addresses the potential for siderophore production in these members through a mass spectrometry-based survey and through bioinformatic analysis of genomic data. This overview of the siderophore potential of symbionts suggests the predomination of catecholate-type siderophores from this community of bacteria. Chapter 4 analyzes how siderophores might be utilized by shipworm symbionts and presents the potential for siderophores to mediate interactions between symbionts. The combined use of cross-streaking experiments and disc-diffusion assays indicate that siderophores have the potential to inhibit the growth of fellow symbionts. This suggests that the endosymbiotic community of shipworms does not work as a cooperative unit in the acquisition of iron.

Introduction

1.1 Shipworms

Shipworms are marine bivalve mollusks of the family Teredinidae. They utilize wood as both a source of shelter and food. They can be found in floating, sunken, or living wood ranging from intertidal zones to depths of 150 m (Distel et al., 2011). Shipworms are globally distributed and can be found in both tropical and temperate waters over a broad range of salinities, with over 65 described species. The name shipworm comes from their ability to cause major destruction to marine wood structures, such as ships. It's estimated that these organisms cause over one billion dollars of damage in the marine environment annually (Distel et al., 2011).

By burrowing into wood, shipworms are able to protect themselves from the external environment. They use relatively small, highly modified shells that cover only the anterior tip of the body to mechanically grind the wood. As it excavates, the shipworm lays down a calcareous lining in the burrow that forms a tube that opens to the external environment. Calcified plates, called pallets, are used to seal off the burrow during unfavorable fluctuations in the environment and are also used for species identification. Reproductive strategies range from internal fertilization and brooding of larvae to broadcast spawning (Distel, 2003).

The wood that is excavated can be digested by the shipworms as a source of food. In shipworm genera such as *Lyrodus* and *Bankia*, the wood enters through the mouth, passes through the stomach, and accumulates in the caecum, the primary site for wood digestion. After digestion, the waste material is transported through the

intestine to the excurrent siphon where it is released into the environment. A simplified schematic of the shipworm anatomy is shown in Figure 1.1.

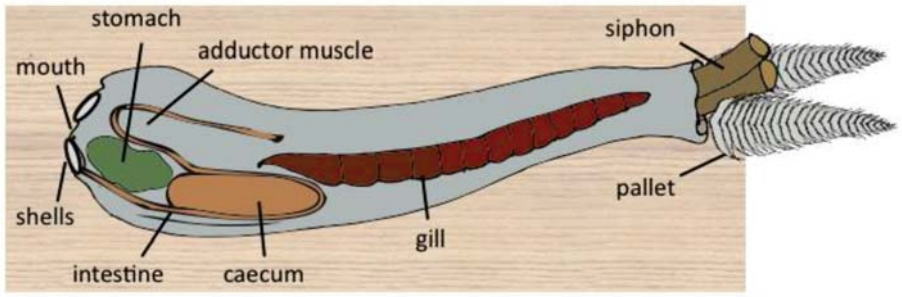

Figure 1.1. Anatomy of the shipworm. Adapted from Meghan Betcher.

A symbiotic community of bacteria that resides in specialized host cells, bacteriocytes, in the gills is believed to aid in the digestion of this wood by the production of cellulolytic enzymes. Woody tissue has two major structural polymers, lignin and cellulose, and is thus referred to as lignocellulose. Cellulose degradation into soluble sugars requires the activity of three types of enzymes: endoglucanases, exoglucanases, and β-glucosidases (Zhang et al., 2006). Most animals that predominately utilize wood as a food source rely on microbially-produced enzymes to solubilize and digest the woody substrate. Therefore it is believed that the endosymbiotic community assists the shipworm in the digestion of lignocellulose.

The assistance by an associated microbial community in the degradation of wood is not unique to the shipworm. A notable example of this type of system is the termite. The termite harbors a mixed community of eukaryotes and prokaryotes in the organ of wood digestion that work together to digest wood (Breznak and Brune,

1994; Ohkuma, 2008). However, in contrast to the termite, shipworms have been reported to contain few microbes in their digestive systems (Greenfield and Lane, 1953; Betcher, 2011). The process of wood digestion in the shipworm is believed to be carried out by cellulolytic enzymes produced by gill-localized symbionts that are excreted into the digestive organs (pers. comm. Jennifer Fung and Dan Distel). The reason for the scarcity of microbes in the digestive tract and the mechanism of cellulolytic enzyme transport are still not fully understood, but these aspects form the basis of a unique symbiont-host system.

Many shipworms are capable of filter feeding (Turner, 1966), but wood serves as the primary source of food for most species (Distel, 2003). A challenge faced when relying on wood as a main nutritional source is the poor source of nitrogen that wood provides. The woody diet must be supplemented with a source of nitrogen. Bacterial symbionts isolated from multiple species of shipworms show the capability to fix nitrogen in laboratory cultures (Distel et al., 2002b). Lechene et al. (2007) directly implicated gill symbionts in fixing nitrogen in host tissue of the shipworm *Lyrodus pedicellatus* by using multiple-isotope imaging mass spectrometry (MIMS) and $^{15}N_2$. Furthermore, Lechene et al. (2007) showed that the newly fixed nitrogen was transferred to host cells and was believed to be metabolized by host cells.

The endosymbiotic community of shipworms, therefore, serves a valuable nutritional role in the shipworm, by the production of cellulolytic enzymes and the fixation of nitrogen. This community has been examined by culture-dependent and independent means (Distel et al., 2002a; Distel et al., 2002b; Luyten et al., 2006), beginning with the cultivation of *Teredinibacter turnerae* (Waterbury et al., 1983).

T. turnerae was once thought to be the only symbiont; however, it has since been shown that *T. turnerae* is part of a group of closely related, yet distinct, gamma proteobacteria inhabiting shipworm gills. The location of these symbionts, as mentioned above, is in bacteriocytes within the gill. The gill spans a large percentage of the length of the shipworm body. Bacteriocytes line the interlamellar junctions and efferent branchial vein within the gills (Figure 1.2) (Distel et al., 1991).

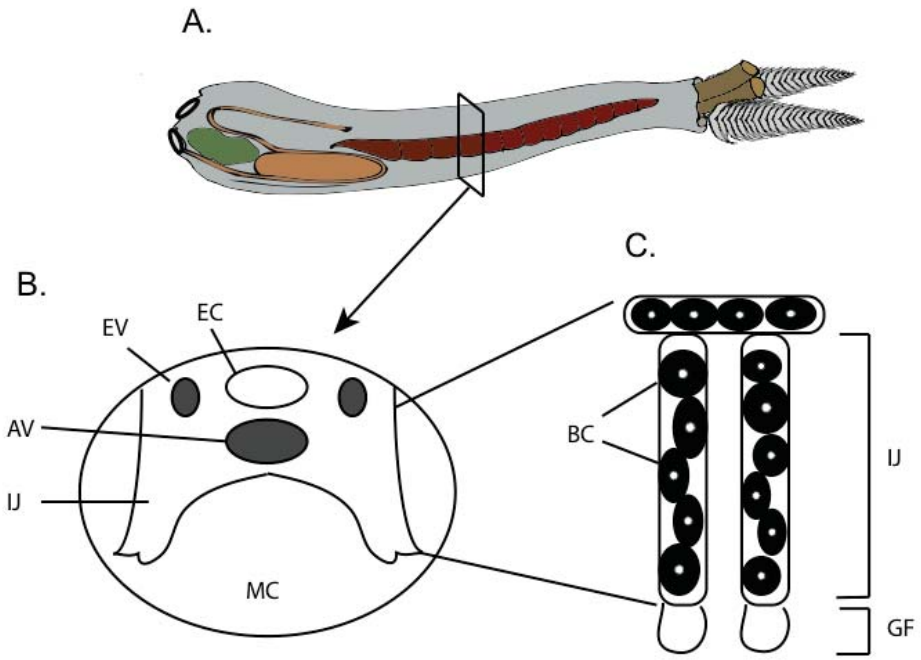

Figure 1.2. Representation of shipworm gill anatomy, adapted from Distel et al. (1991) and Turner (1966). (A) Shipworm body. (B) Cross-section of gills and associated blood vessels. MC, mantle cavity; IJ, intralamellar junction; AV, afferent branchial vein; EV, effecrent branchial vein; EC, epibranchial cavity. (C) Tangential sections of gills. GF, gill filament, BC; bacteriocyte.

1.2 Nitrogen fixation

Nitrogen is an essential element for all organisms because it is a component of many biomolecules, including proteins and nucleic acids. The largest pool of nitrogen in the biosphere is atmospheric dinitrogen (N_2), but this form is unusable by

most organisms due to its inert nature. Biological nitrogen fixation is the enzymatic conversion of atmospheric N_2 gas into ammonia, represented by the following equation:

$$N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i.$$

The two ammonia molecules are subsequently converted into other usable forms of nitrogen.

Nitrogen fixation is catalyzed by the nitrogenase enzyme. The most common form of nitrogenase that exists in all diazotrophs is the molybdenum nitrogenase, but some organisms also carry a vanadium and/or an iron-only nitrogenase (Bishop and Joerger, 1990). The molybdenum nitrogenase enzyme contains a homodimeric component (NifH) that contains an Fe_4-S_4 metallocluster and a heterotetrameric component (NifDK) that contains a Mo-Fe_7-S_9 cofactor and a Fe_8-S_7 P cluster (Rubio and Ludden, 2008). Thus, iron is considered a critical nutrient for the process of nitrogen fixation.

1.3 Iron limitation

Work with the diazotrophic cyanobacterium *Trichodesmium* estimates that the requirements for iron under diazotrophic growth is five times higher than under ammonium-supported growth (Kustka et al., 2003). Studies looking at oceanic nitrogen fixation suggest that iron is the limiting nutrient in the distribution of this process (Mills et al., 2004; Moore et al., 2009). The insolubility of Fe(III) at physiological pH and aerobic conditions severely limits its bioavailability. In addition, 99% of the Fe(III) in the ocean is bound by uncharacterized organic ligands

(Rue and Bruland, 1995; Wu and Luther, 1995). Similarly, iron is also a limiting nutrient in the host, where it is often tightly bound to host proteins such as ferritin, transferrin, and heme containing proteins (Ratledge and Dover, 2000; Chu et al., 2010). In the mammalian host, it's been estimated that the iron requirements of a bacterium for various metabolic processes are ~10^{12} to 10^{18} fold higher than the free iron that is available in the host (Andrews et al., 2003; Raymond et al., 2003). The ability to acquire iron by associated or infecting bacteria is critical in establishing a presence within the host.

Several studies have identified iron acquisition systems as integral components that allow bacteria to colonize their respective hosts. Studies with the fish pathogen *Vibrio anguillarum* have shown that virulence is dependent upon the presence of a plasmid-encoded iron-acquisition system (Crosa, 1980; Crosa et al., 1980; Wolf and Crosa, 1986). In the symbiosis between the luminescent bacterium *Vibrio fischeri* and the squid *Euprymna scolopes*, colonization by a mutant *V. fischeri* strain, which was partially deficient in its ability to acquire iron, failed to persist (Graf and Ruby, 2000). The addition of $FeCl_3$ relieved this defect. Similarly, the addition of $FeCl_3$ relieved the defect found in an iron uptake mutant of the nematode symbiont *Photorhabdus.* The mutant *Photorhabdus* was found to be unable to support the growth and development of its host (Watson et al., 2005). These studies highlight the need to study the iron acquisition systems of symbionts such as the gill endosymbionts of shipworms. Their implicated role in nitrogen fixation suggests that they may have an even higher demand for iron in the host.

1.4 Siderophores

In response to iron limitation, many bacteria and some fungi produce low-molecular weight compounds with a high affinity for Fe(III) called siderophores. Hundreds of siderophores have been characterized, the majority being from the terrestrial environment (Raymond et al., 1984; Winkelmann, 2002; Sandy and Butler, 2009). The three major types of functional binding groups are catechols, hydroxamic acids, and α-hydroxycarboxylic acids (Figure 1.3).

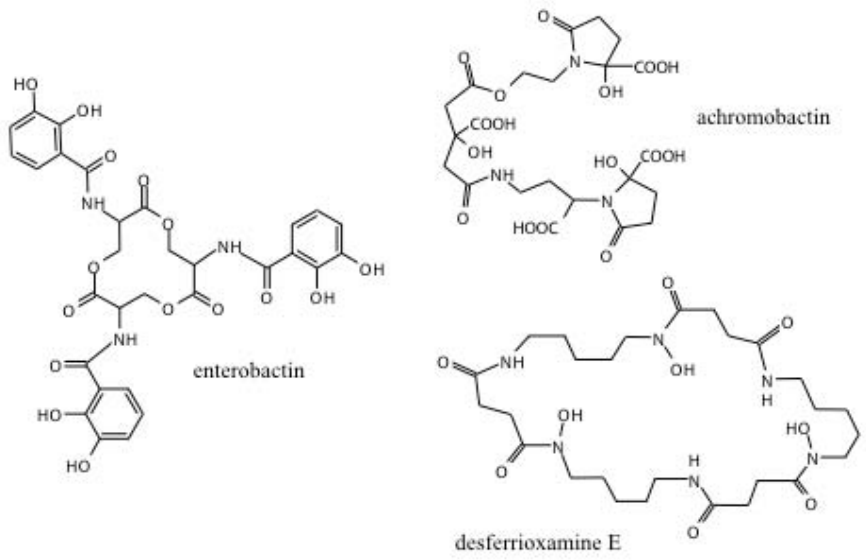

Figure 1.3. Representative siderophore structures showing the three common iron binding motifs: the catecholate siderophore enterobactin, the hydroxamate siderophore desferrioxamine E, the α-hydroxycarboxylate siderophore achromobactin.

The number of siderophore structures from marine microorganisms is relatively few compared to their terrestrial counterparts. However, from the studies available, there are two major themes concerning marine-derived siderophores (Butler, 2005; Vraspir and Butler, 2009). The first being the reporting of a class of amphiphilic siderophores that contain a peptidic iron-binding headgroup appended

with a series of fatty acids (Martinez et al., 2000; Martinez et al., 2003; Ito and Butler, 2005; Martin et al., 2006). The other major feature of marine siderophores is the presence of α-hydroxycarboxylic acids moieties that, when coordinated to Fe(III), are photoreactive (Barbeau et al., 2001; Barbeau et al., 2002; Barbeau et al., 2003; Küpper et al., 2006).

The incorporation of ferric-siderophore complexes is via specific outer membrane receptors in a process driven by the cytosolic membrane potential and mediated by the energy-transducing TonB-ExbB-ExbD system (Andrews et al., 2003). Periplasmic binding proteins shuttle the ferric-siderophore complex to cytosolic membrane ABC-transporters, which deliver the ferric-siderophore complex to the cytoplasm. There iron is released either by reduction of the ferric iron to ferrous iron (Hallé and Meyer, 1992), or by cleavage of the siderophore (Brickman and McIntosh, 1992).

1.5 Coordination of Fe(III) by siderophores

Siderophores exhibit a wide range of affinity for Fe(III). The standard convention for expressing the overall equilibria of metal-ligand stability constants is β_{mlh} for the reaction $m\mathrm{M}+l\mathrm{L}+h\mathrm{H}=\mathrm{M}_m\mathrm{L}_l\mathrm{H}_h$, where M is metal, L is ligand, and H is proton. One of the weakest siderophores described, aerobactin, containing both hydroxamate and α-hydroxycarboxylate groups, has a stability constant (β_{110}) of $10^{22.5}$ (Harris et al., 1979), while the strongest siderophore, enterobactin, containing three catecholate groups, has a stability constant of 10^{49} (Loomis and Raymond, 1991). Another siderophore, petrobactin, that contains two catecholate and an α-

hydroxycarboxylate binding group, and has a stability constant of 10^{43} (Zhang et al., 2009). However, these stability constants do not account for the competitive reaction between protons and the metal to the binding groups of the siderophore, necessitating the need to take into account the pK_a of donor groups when comparing siderophores' effectiveness at complexing Fe(III). A value that allows for a more convenient comparison of binding affinities between siderophores is pM. The pM value measures the logarithmic concentration of free iron for a defined set of experimental conditions: pH 7.4, Fe(III) concentration of 10^{-6} M, and ligand concentration of 10^{-5} M. The calculated pM values of enterobactin and petrobactin are 35.5 (Harris et al., 1979) and 23.1 (Zhang et al., 2009), respectively. This comparison between enterobactin and petrobactin highlights the effect of pK_a of donor groups and the high affinity for Fe(III) of triscatecholate siderophore at physiological pH.

Enterobactin was originally isolated from the enteric bacteria *Escherichia coli* (O'Brien and Gibson, 1970) and *Salmonella typhimurium* (Pollack and Neilands, 1970) and is composed of three units of 2,3-dihydroxybenzoate that are each amide-linked to a trilactone backbone of L-serine. In addition to enterobactin, other triscatechol amide siderophores have been reported. The glucosylated derivative of enterobactin, salmochelin, was reported from *Salmonella enterica* (Hantke et al., 2003; Bister et al. 2004). *Bacillus* spp. produce bacillibactin, a trilactone of 2,3-dihydroxybenzoyl-glycyl-L-threonine (May et al., 2001). The calculate stability constant of bacillibactin,$10^{47.6}$ (Dertz et al., 2006), is comparable to enterobactin but the glycine spacer diminishes the overall stability of the ferric siderophore complex. Until recently, these were the only known triscatechol amide siderophores. But

recent studies have reported additional members of this family of compounds. A trilactone of 2,3-dihydroxybenzoyl-arginyl-threonine was proposed to be produced by strains of Streptomyces (Patzer and Braun, 2009). A linear trimer of 2,3-dihydroxybenzoyl-D-arginyl-L-serine was described from the open ocean bacterium *Vibrio* sp. DS40M4, trivanchrobactin (Sandy et al., 2010). The plant pathogen *Dickeya chrysanthemi* was reported to produce the trilactone of 2,3-dihydroxybenzoyl-D-lysyl-L-serine, cyclic trichrysobactin (Sandy and Butler, 2011).

1.6 Siderophore biosynthesis

Many siderophores are produced by nonribosomal peptide synthetases (NRPS). NRPS's are multimodular enzymes that produce a broad range of complex peptide secondary metabolites. In addition to the 20 proteinogenic amino acids, NRPS's can incorporate a large number of nonproteinogenic amino acids. The products of NRPS's have a wide range of biological activity for use in antibiotic and therapeutic applications (Fischbach and Walsh, 2006; Mootz et al., 2002). Examples include the antibiotic glycopeptides of the vancomycin family (van Wageningen et al., 1998) and the precursors for the penicillin and cephalosporin families of antibiotics (Martín, 1998) (Figure 1.4).

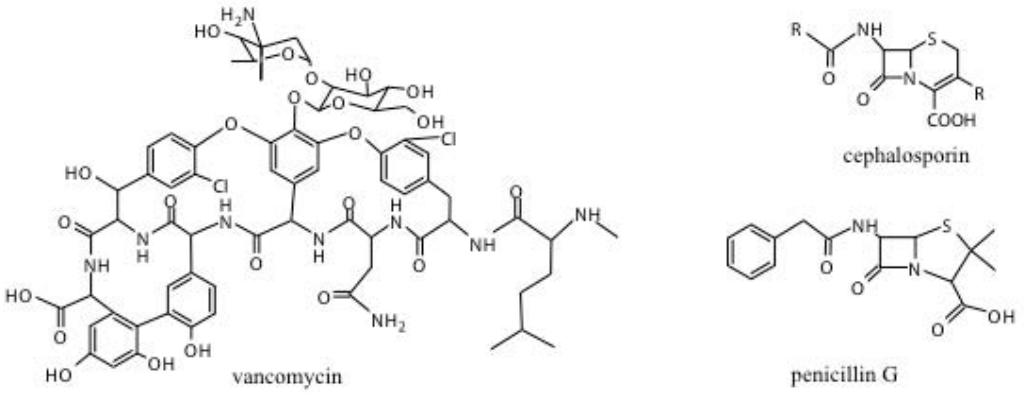

Figure 1.4. Antibiotic compounds produced by NRPS's.

NRPSs are composed of modules, with each module responsible for incorporating one amino acid into the peptide product. There are three core domains that make up a minimal module. An adenylation domain (A domain) selects and activates an amino acid substrate by converting it to an aminoacyl adenylate via an attack by ATP. The activated aminoacyl adenylate is transferred to the thiol moiety of a 4'-phosphopantetheine (P-pant) attached to a peptidyl carrier protein (PCP domain) to form an aminoacyl-thioester (aminoacyl-S-PCP). The P-pant group of each PCP domain is posttranslationally introduced onto the hydroxy group of a conserved serine (Ser) residue within the PCP domain by a P-pant transferase (Lambalot et al., 1996). The PCP domain serves to transport the aminoacyl and peptidyl chains to the various catalytic domains of the NRPS. Condensation domains (C domain) catalyze the formation of the peptide bond of the upstream peptidyl-S-PCP to the downstream aminoacyl-S-PCP, thereby elongating the peptidyl chain by one aminoacyl residue. Chain termination and peptide release is usually carried out by a C-terminus thioesterase domain (TE domain). The peptide chain is transferred

from the terminal PCP domain to a serine residue in the TE domain to form an acyl-O-TE intermediate. The final peptide product can be released either by a hydrolytic reaction or an intramolecular cyclization reaction. Alternatively, a C-terminal reductase domain (Re domain) can catalyze peptide release by a reduction of the thioester, resulting in the release of an aldehyde product.

Additional NRPS tailoring domains exist that modify substrates. Many NRPS's contain epimerase domains (E domain) that generate D-amino acids. The domain organization in such modules would be C_n-A_n-PCP_n-E_n-C_{n+1}-A_{n+1}-PCP_{n+1}, where the E_n domain converts the L-amino acid activated by the A_n domain to its D configuration and the C_{n+1} domain catalyzes the peptide bond between the D-amino acid selected by the A_n domain to the L-amino acid selected by the A_{n+1} domain. These C domains have thus been designated as $^{D}C_{L}$ domains, for D-donor and L-acceptor. This would be in contrast to $^{L}C_{L}$ domains where there was no E domain present. Recently, a new type of C domain was discovered that has dual epimerization and condensation activity (Dual E/C domain) (Balibar et al., 2005), which can be viewed as $^{D}C_{L}$ domains with epimerase activity. The Dual E/C domain catalyzes the epimerization of the upstream L-amino acid to a D configuration and subsequently catalyzes the condensation of the upstream and downstream resides. Another type of C domain is the heterocyclization domain (Cy domain). These domains catalyze the peptide bond formation and the subsequent cyclization of cysteine (Cys), Ser, or threonine (Thr) residue, resulting in a heterocyclic five-membered ring. Another tailoring domain is the methyltransferase domain (MT

domain), which catalyzes the transfer of the methyl group from *S*-adenosylmethionine to the amino group of an aminoacyl-S-PCP intermediate in a C-A-MT-PCP module.

In addition to NRPS-dependent routes of siderophore biosynthesis, NRPS-independent routes of siderophore biosynthesis exist, utilizing alternating dicarboxylic acid and diamine or amino alcohol building blocks linked by amide or ester linkages (Challis, 2005). The most notable example of this type of siderophore is aerobactin, which utilizes four genes, *iucABCD*, for biosynthesis (de Lorenzo et al., 1986). IucD begins catalyzing the N6-hydroxylation of lysine, to form N6-hydroyllysine. Next, IucB catalyzes the N6-acetylation of N-6-hydroylysine with acetyl CoA to form N6-acetyl-N6-hydroxylysine. Then, IucA catalyzes the condensation of N6-acetyl-N6-hydroxylysine with one of the prochiral carboxyl groups of citric acid. This product is then condensed with another unit of N6-acetyle-N6-hydroxylysine by IucC to form aerobactin.

While it is important for bacteria to secure adequate supplies of iron required for growth, it is equally important to maintain intracellular iron concentrations at safe levels. Excess intracellular iron concentrations can have toxic effects by catalyzing the formation of reactive oxygen species via Fenton reactions. Therefore, iron uptake genes need to be carefully regulated. This regulation is mediated by the ferric uptake regulator (Fur) protein (Andrews et al., 2003; Escolar et al., 1999). The Fur protein is a homodimer made up of 17 kDa subunits that acts as a transcriptional regulator of iron-regulated promoters. Fur binds ferrous iron and undergoes a conformational change that allows it to bind to target DNA sequences, called Fur boxes. Upon binding, transcription of the operon is inhibited. When iron in the cell is scarce, Fur

is not bound to iron and, thus, the promoter is free allowing access to RNA polymerase for gene transcription. Studies with *Escherichia coli* originally proposed that Fur proteins recognize a 19-bp inverted repeat sequence, GATAATGATAATCATTATC (de Lorenzo et al., 1987). However, more recently, Baichoo and Helmann (2002) proposed a revised view of the Fur box, suggesting that the Fur box is composed of two overlapping inverted repeats, each consisting of a 7-1-7 motif and each capable of binding a Fur protein, with a consensus sequence of TGATAAT-ATTATCA.

1.7 Overview of the current study

The shipworm presents a unique system in which to study microbial community dynamics. The endosymbiotic community is integral in supporting the nutrition of the shipworm host. However, the interactions between symbionts in carrying out these processes are not well understood. The goal of this work is to investigate how siderophores influence the microbial community dynamic in the shipworm. This will be carried out by examining the biosyntheses, structures, and mediation activities of siderophores by bacterial endosymbionts of the shipworm. First, the model symbiont *Teredinibacter turnerae* will be presented along with the structure and biosynthesis of its siderophore. Next, a profile of the siderophore capacity of the shipworm symbionts as a whole will be presented. Finally, the potential for siderophores to mediate interactions between symbionts will be presented and discussed.

This work is in conjunction with the work of the Philippine Mollusk Symbiont International Conservation Biodiversity Group (PMS-ICBG) and the laboratory of Dan Distel. The PMS-ICBG is part of an NSF-funded program that addresses the interdependent issues of drug discovery, biodiversity conservation, and sustainable economic growth. These two entities have provided valuable resources in the form of genomic data and bacterial strains.

2. The siderophore from model symbiont *Teredinibacter turnerae* T7901

2.1 Introduction

The shipworm symbiont *Teredinibacter turnerae* has been found in numerous genera and species of teredinid bivalves from across the globe (Distel et al., 2002b). *T. turnerae* was originally cultivated by Waterbury et al. (1983) and was formally described by Distel et al. (2002b). Although it has been estimated that *T. turnerae* makes up only about 10% of the bacterial community in the shipworm *Lyrodus pedicellatus* (Luyten et al., 2006), it is believed to be an overall important member of the endosymbiotic community due to its presence in a total of nine genera and 24 species of teredinid bivalves (Distel et al., 2002b).

The genome of *T. turnerae* T7901 was recently sequenced (Yang et al., 2009). Although *T. turnerae* is an intracellular symbiont, its genome does not show the typical modifications experienced by obligate endosymbionts, most notably reduced genome size, higher A+T content, and loss of DNA repair and transcriptional regulatory genes (Moya et al., 2008). Its genome size is 5.1 Mb, with a 50.8% G+C content, and includes genes involved in almost all core metabolic functions, including DNA repair (Yang et al., 2009). These observations suggest either *T. turnerae* is a facultative symbiont, or is part of a recently established symbiosis. To date, *T. turnerae* has not been detected outside of the shipworm host.

The genome of *T. turnerae* T7901, as well as extensive laboratory data, also reveals its ability to fix nitrogen and produce cellulolytic enzymes, both important functions that nutritionally support the host (Waterbury et al., 1983; Distel et al., 2002b; Distel, 2003; Lechene et al., 2007; Yang et al., 2009). The *nif* operon contains

a complete set of genes for nitrogen fixation. The suite of cellulases present in the genome suggests that *T. turnerae* is highly specialized at degrading woody plant material, such as cellulose, xylan, mannan, galactorhamnan, and pectin (Yang et al., 2009).

Analysis of *T. turnerae*'s genome also reveals nine gene clusters predicted to code for secondary metabolites, constituting nearly 7% of its genome. The field of natural products is constantly searching for new bioactive compounds with novel activity, particularly from symbionts from marine invertebrates (Piel, 2006). The production of these compounds may play a role in the symbiosis, either in antimicrobial activity against unwanted microbes or chemical defense of the shipworm and its progeny.

One of the nine gene clusters of the secondary metabolome putatively codes for a siderophore biosynthetic gene cluster (Yang et al., 2009). Previous work with this gene cluster examined the expression of this cluster under varying conditions and showed upregulation in iron-limiting conditions (Fishman, 2010), supporting the role of this gene cluster in iron acquisition.

This work describes a gene cluster responsible for the biosynthesis of a novel siderophore, 2,3-dihydroxybenzoyl-L-ornithyl-L-serine, named turnerbactin, by *Teredinibacter turnerae* T7901. The biosynthetic genes involved in turnerbactin production are herein referred to as *tnb*, for *t*ur*n*erbactin *b*iosynthesis. Turnerbactin is structurally characterized and a model for its biosynthesis is proposed. Turnerbactin represents the first described siderophore from a shipworm symbiont.

2.2 Methods

2.2.1 Strain information

Teredinibacter turnerae T7901 was originally isolated from the shipworm *Bankia gouldi* from a laboratory-reared shipworm culture maintained at the Woods Hole Oceanographic Institute, MA. The original shipworm specimen was from Duke University Marine Lab, Beaufort, NC (Distel et al., 2002b).

2.2.2 Culture conditions

T. turnerae T7901 was grown in a low-iron, modified version of shipworm basal medium (SBM) (Distel et al., 2002b; Waterbury et al., 1983) containing 750 ml artificial seawater (Kester et al., 1967), 250 ml distilled water, 0.1 mM KH_2PO_4, 0.094 mM Na_2CO_3, 0.01 mM $Na_2MoO_4.2H_2O$, 0.5% (w:v) sucrose, 5 mM NH_4Cl, 20 mM HEPES buffer (pH 8.0), 0.1 uM EDTA-chelated ferric iron (Sigma), and 1 ml A5+Co trace metal mix (Rippka et al., 1979). Cultures were grown in 2.8 L Fernbach flasks with 2 L culture medium at 30° C on an orbital shaker at 110 rpm. A pre-culture of *T. turnerae* was prepared by inoculating an iron-replete (10 uM EDTA-chelated ferric iron) SBM liquid culture with a single colony from an iron-replete SBM agar plate (1.5% agar, w:v). The 2 L cultures were inoculated with 1 mL of overnight-grown pre-culture. Cultures were grown for 2 days, when chrome azurol sulfonate (CAS) (Schwyn and Neiland, 1987) activity of culture supernatants reached maximum activity. All growth measurements were recorded on a SpectraMax M2 Multidetection Reader (Molecular Devices) with absorbance measured at 600 nm.

2.2.3 Compound isolation

Culture supernatant was harvested by centrifuging cultures at 10000 rpm for 25 min. Decanted supernatant was placed in new 2.8 L Fernbach flasks, ~20 g/L Diaion HP20 resin was added to the flasks, and the flasks were then incubated at 4° C on an orbital shaker at 110 rpm for 4 hrs. The mixture was poured into a column allowing the spent supernatant to flow through. The HP20 resin was washed with MilliQ water, followed by 25%, 50%, and finally 100% isopropanol. The siderophores were eluted with the 25% isopropanol fraction as detected by the CAS assay. CAS activity was not detected in the 50% or 100% fraction. The 25% isopropanol fraction was concentrated by rotary evaporation *in vacuo*. The siderophores were purified by reverse-phase high–performance liquid chromatography (RP-HPLC) on a semi-preparative C_{18} column (10 mm internal diameter x 25 cm length, 5 μm particle size, Ascentis) using water with 0.05% trifluoroacetic acid (TFA) and acetonitrile (ACN) with 0.05% TFA as mobile phases at a flow-rate of 3 mL/min. The programmed run consisted of a gradient run of 10% ACN to 15% ACN from 0-10 min, followed by an isocratic step of 15% ACN from 10-15 min, then a gradient from 15% to 25% ACN from 15-35 min, then a final gradient from 25% to 100% methanol from 35-40 min. The eluent was continuously monitored at 215 nm and 320 nm. Fractions were collected manually and immediately tested for CAS and Arnow (Arnow, 1937) activity and then concentrated under vacuum. Collected fractions were further purified on an analytical RP-Amide-C_{16} column (4.6 mm internal diameter x 25 cm length, 5 μm particle size, Ascentis) using the same method as above at a flow-rate of 1 mL/min. Pure siderophores were

lyophilized. The approximate yields of siderophores from a 4 L culture are 1.2 mg turnerbactin, 2.1 mg diturnerbactin, 2.7 mg triturnerbactin, 2.4 mg dehydrated diturnerbactin, and 1.2 mg dehydrated triturnerbactin.

Mass spectra were obtained on a ThermoElectron LTQ-Orbitrap high resolution mass spectrometer as part of the Portland State University Bioanalytical Mass Spectrometry Facility. Samples were dissolved in 50% methanol and MS analysis was taken in positive mode using electrospray ionization (ESI). All 1-D and 2-D NMR experiments were carried out on a Bruker Avance II Ultrashield Plus 800 MHz instrument with a cryoprobe in d_4-methanol (CD_3OH). The NMR instrument is part of the NMR facility at the University of California Santa Barbara, made available courtesy of Prof. Alison Butler with assistance from Moriah Sandy.

2.2.4 Amino acid analysis

A dried preparation of purified triturnerbactin (~1 mg) was hydrolyzed in 200 μL 6 M HCl for 17 hours at 110° C. The solution was brought to room temperature and evaporated to dryness. The dried, hydrolyzed sample was redissolved in 100 μL H_2O, to which 200 μL of a 1% (w:v) solution of Marfey's reagent (1-fluoro-2-4-dinitrophenyl-5-L-alanine amide) (Marfey and Ottesen, 1984) in acetone and 40 μL 1 M $NaHCO_3$ were added. The reaction was heated for 1 hour at 40° C, after which 20 μL 2 M HCl was added to stop the reaction. The derivatized sample was analyzed by HPLC on an analytical YMC ODS-AQ C_{18} column (4.6 mm, i.d. x 250 mm L, Waters Corp.) using a linear gradient from 90% triethylamine phosphate (TEAP) (50 mM; pH 3.0)/10% CH_3CN to 60% TEAP/40% CH_3CN over 45 min. The eluent was

continuously monitored on a Waters UV-visible detector at 340 nm. The derivatized sample was compared to chiral amino acid standards prepared the same way.

2.2.5 Chemical extraction of shipworms

A laboratory culture of *Lyrodus pedicellatus* was obtained from Dan Distel at the Ocean Genome Legacy and was maintained at room temperature in an aerated aquarium until further processed. *L. pedicellatus* specimens were removed from wood using pliers, slowly peeling away layers of wood until shipworms were exposed, careful to not damage the shipworms. Six *L. pedicellatus* shipworms were removed from the wood, rinsed in filter-seawater, pooled, and lyophilized. The resulting dry weight was 85.4 mg. The tissue was then homogenized with a plastic pestle. The tissue was extracted three times with three volumes of methanol on a rotary mixer for one hour. The methanolic extract was separated from the tissue by centrifugation at 15,000 rpm for 3 min and subsequently dried. Approximately 1 mg of the crude methanolic extract was dissolved in water and injected onto a ThermoElectron LTQ-Orbitrap high resolution mass spectrometer with a dedicated Accela HPLC system using a C_{18} column (2.1 mm internal diameter x 10 cm length, 3 μm particle size, Ascentis). Mass spectra were recorded in positive mode using ESI. The following gradient run was used at a flow rate of 200 ul/min: 10% methanol in water to 50% methanol over 30 min, followed by an increase to 100% methanol over 5 min. Both solvents contained 0.1% formic acid. Pure turnerbactin (approximately 10 μg) was used as a standard using the same method. The *L. pedicellatus* extract

was analyzed first, followed by washing of the column, a blank injection, and then the pure turnerbactin standard.

2.2.6 Bioinformatic analysis of NRPS domains

Module identification and domain organization of the NRPS was carried out by the online tools NRPS-PKS (Ansari et al., 2004) and the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Additional analysis utilized the software package HMMer (Eddy, 1998) available from the website http://hmmer.org/.

Adenylation domain: The specificity-conferring code of *tnbF* A domains were determined using three different online tools: NRPS-PKS (Ansari et al., 2004), NRPSpredictor2 (Rottig et al., 2011), and the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Phylogenetic analysis of the specificity-conferring codes of numerous A domains was also carried out. Protein sequences of A domains in *tnbF* were retrieved using the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Reference A domain protein sequences were compiled from a list of A domains with annotated specificity supplied by Rausch et al. (2005). A multiple sequence alignment was constructed using MUSCLE (Edgar, 2004). After extracting approximately a 100 amino acid stretch between core motifs A4 and A5, the alignment was manually edited, guided by the structural anchors of this region identified by Stachelhaus et al. (1999) and Challis et al. (2000). The alignment was manipulated and edited in Geneious v5.4 (Drummond et al., 2011). The eight residue alignment was exported out of Geneious and into MEGA5 (Tamura et al., 2011). Phylogenetic trees were inferred using the unweighted pair group method arithmetic

averages (UPGMA) (Sneath and Sokal, 1973) with the JTT (Jones et al., 1992) amino acid substitution matrix. The rate variation among sites was modeled with a gamma distribution.

Condensation domain: Protein sequences of condensation domains in *tnbF* were retrieved using the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Aligned protein sequences of condensation domains were downloaded from Rausch et al. (2007). The condensation domain from *tnbF* was added to the downloaded condensation domain alignment using the multiple sequence alignment program MUSCLE. The alignment was edited in Geneious v5.4. A maximum likelihood (ML) tree of condensation domains was reconstructed using RAxML 7.2.7 (Stamatakis, 2006). The amino acid substitution matrix used in this analysis was the JTT matrix (Jones et al., 1992), with the Γ model of rate heterogeneity. RAxML's rapid bootstrap was performed with 100 replicates and the best scoring ML tree was saved. This analysis was done using the facilities made available by the CIPRES portal in San Diego, CA (Miller et al., 2010).

TE domain: The sequence of the TE domain was aligned with other annotated NRPS TE domains using MUSCLE. The alignment was edited in Geneious v5.4. The curation of the alignment was guided by the alignments presented in the work by Bruner et al. (2002) and Samel et al. (2006), who reported the structures of SrfTE and FenTE, respectively.

2.2.7 Construction of *tnbF* plasmid insertion disruption mutant

For construction of the plasmids used for gene disruption, a region targeting the first C domain within *tnbF* was amplified with specific PCR primers (Table 2.1) and the high-fidelity Phusion DNA polymerase (Finnzymes). The addition of 3' A-overhangs for TA cloning was carried out by DyNAzyme II DNA polymerase (Finnzymes) following the manufacturer's protocol. The amplicon was purified using the QIAquick PCR Purification Kit (Qiagen) and then cloned into pCR2.1 (Invitrogen). Plasmid DNA was isolated and double digested with XhoI and SacI. The resulting fragment was gel purified with the GENECLEAN II Kit (MP Biomedicals) and then ligated with the Quick Ligation Kit (NEB) into suicide vector pDM4, which had been double digested with the same restriction enzymes, resulting in plasmid pDMtnbF. The plasmid pDMtnbF was transformed into *E. coli* S17-1 λpir. This plasmid was then conjugated into *T. turnerae* T7901, and plasmid cointegrates were selected on SBM plates with Sigmacell cellulose Type 101 (Sigma) as the sole carbon source and 10 μg/ml chlorampenicol (Cm). The location of integration by pDMtnbF into the chromosome of T7901 was confirmed by PCR and DNA sequencing.

Table 2.1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Description	Reference
Teredinibacter turnerae strains		
T7901	wt	Distel et al., 2002b
TtAH03	*tnbF*::Cm^r	This study
Escherichia coli strains		
TOP10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^r) *end*A1 *nup*G	Invitrogen
S17-1 λpir	*thi pro hsdR hsdM*$^+$ *recA* RP4-2-Tc::Mu-Km::Tn7λpir	Simon et al., 1983
Plasmids		
pCR2.1	T-vector, Km^r, Amp^r	Invitrogen
pDM4	Suicide plasmid, *sacB* gene, R6K origin, Cm^r	Milton et al., 1996
pDMtnbF	portion of tnbF gene cloned into pDM4	This study

2.3 Results

2.3.1 Biosynthetic gene cluster

Annotation of the genome of *Teredinibacter turnerae* T7901 revealed a cluster of genes with similarity to siderophore biosynthesis and iron transport genes (Yang et al., 2009). The organization of the gene cluster is shown in Figure 2.1. Proteins with closest similarity to the gene products of this cluster are shown in Table 2.2 and suggest the production of an enterobactin-like catecholate siderophore. The genes *tnbCEBA* are homologous to the *entCEBA* genes, responsible for the biosynthesis and activation of 2,3-dihydroxybenzoate (DHB) via the shikimate pathway (Walsh et al., 1990). The product of gene *entC* isomerizes chorismate into isochorismate, then the N-terminal portion of *entB* hydrolyzes isochorismate into 2,3-dihydro-DHB, which is oxidized into DHB by *entA*. Then *entE* activates and transfers DHB to the aryl carrier C-terminal portion of *entB*. The gene *tnbF* codes for a 2399 amino acid, two module NRPS. Each module contains a C, A, and PCP

domain, followed by a C-terminus TE domain (Figure 2.1). A phosphopantetheinyl (P-pant) transferase, which is required for NRPS biosynthesis, was not found at this locus. The genome of *T. turnerae* T7901 codes for two P-pant transferases, TERTU_1510 and TERTU_4652. Due to its homology to the *entD* gene of enterobactin biosynthesis, it is believed that TERTU_1510 is the P-pant transferase that posttranslationally attaches a P-pant moiety to a conserved serine residue of the apo-PCP domains of *tnbF*. Further experiments are necessary to confirm this.

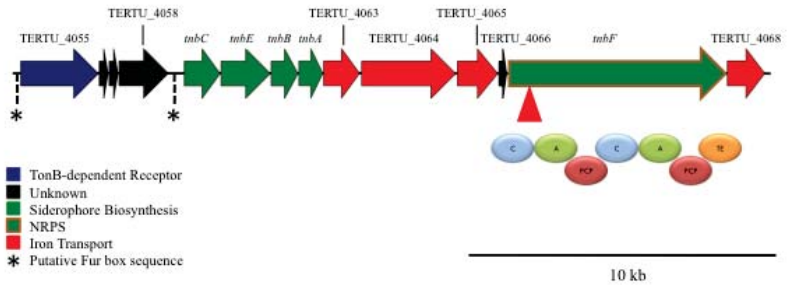

Figure 2.1. Organization of the turnerbactin biosynthetic gene cluster. Shown below *tnbF* is its predicted NRPS domain organization. The red arrow below *tnbF* indicates the insertion site of the chloramphenicol resistance marker for targeted mutagenesis (see section 2.3.4)

Table 2.2. Proteins with similarity to the products of the turnerbactin biosynthetic gene cluster. S: similarity, I: identity.

Locus	Protein size	Similarity	S/I (%/%)
TERTU_4055	648	TonB-dependent receptor (*Acinetobacter calcoaceticus*)	60/40
TERTU_4056	87	Methyl-accepting chemotaxis sensory transducer (*Shewanella amazonensis*)	50/37
TERTU_4057	100	Hypothetical protein *(Shewanella* sp. ANA-3)	70/55
TERTU_4058	550	Response regulator receiver protein *(Shewanella* sp. ANA-3)	63/47
tnbC	394	DhbC (*Brucella suis*)	65/50
tnbE	538	EntE (*Brucella ceti*)	75/62
tnbB	292	DhbB (*Photobacterium profundum*)	72/53
tnbA	255	DhbA (*Aeromonas hydrophila*)	71/57
TERTU_4063	389	Secretion protein HlyD (*Shewanella denitrificans*)	65/44
TERTU_4064	1047	Cation/multidrug efflux pump (*Hahella chejuensis*)	75/60
TERTU_4065	445	Enterobactin esterase (*Dickeya dadantii*)	50/38
TERTU_4066	85	MbtH domain protein (*Dickeya dadantii*)	72/55
tnbF	2399	DhbF (*Acinetobacter baumannii*)	62/45
TERTU_4068	417	Enterobactin exporter entS *(Acinetobacter* sp. RUH2624)	72/52

TERTU_4066 codes for an MbtH-like protein (Quadri et al., 1998), a protein of unknown function that is found in many but not all NRPS gene clusters. TERTU_4068 codes for a homologue of *entS*, a 12 trans-membrane domain-containing efflux pump belonging to the Major Facilitator Superfamily (MFS) shown to excrete enterobactin in Escherichia coli (Furrer et al., 2002). TERTU_4055 codes for a putative TonB-dependent receptor, required for the recognition and uptake of the ferric siderophore complex. TERTU_4065 has close homology to the enterobactin esterase, *fes*. In *E. coli*, once ferric enterobactin has entered the cell, iron is removed from the siderophore by hydrolyzing the ester bonds of the siderophore's backbone with the *fes* gene product.

This gene cluster also shows two putative Fur box sequences (Figure 2.1), suggesting transcriptional regulation by one of the three Fur homologs found in the

genome, TERTU_0053, TERTU_3299, and TERTU_3389. These sequences are compared to the bacillibactin biosynthetic operon of *Bacillus subtilis* in Table 2.3, whose Fur box matches exactly to the classic Fur box sequence proposed by de Lorenzo et al. (1987). The Fur box sequences of turnerbactin's biosynthetic gene cluster show high similarity with the two overlapping 7-1-7 motifs proposed by Baichoo and Helmann (2002). This suggests that the turnerbactin biosynthetic gene cluster is composed of two iron-regulated operons: the first operon consisting of the genes TERTU_4055-4058, and the second operon consisting of the genes *tnbC*-TERTU_4068.

Table 2.3. Putative Fur box sequences of turnerbactin's biosynthetic gene cluster compared to the bacillibactin biosynthetic operon of *Bacillus subtilis*. Shaded bp's match the revised view of Fur box sequences consisting of two overlapping 7-1-7 motifs proposed by Baichoo and Helmann (2002).

Operon	**7-1-7 consensus sites** TGATAAT-ATTATCA TGATAAT-ATTATCA (L) TGATAAT-ATTATCA (R)
dhbABCKF	ATTGATAATGATAATCATTATCAATAGATTG
TERTU 4055-58	AATGATAATGCGAATAATTATCATATAGATT
tnbC-TERTU 4068	CACAATAATGCGAATCATTCTCATTATTGAT

2.3.2 NRPS domain specificities

The gene *tnbF* codes for a two module NRPS. Analysis of the catalytic domains reveals their respective specificity and allows for a prediction of the compound.

2.3.2.1 A domain. The crystal structure of the phenylalanine activating A domain of gramicidin synthetase A (GrsA) by Conti et al. (1997) identified 10 residue positions critical in substrate binding. Stachelhaus et al. (1999) and Challis et al.

(2000) determined that these 10 residues correspond to similar residues in other A domains and that these residues would form a "specificity-conferring code". They surmised that this code could be extracted from uncharacterized A domains and could be compared to known A domains to infer specificity.

Three online tools were used to identify the specificity-conferring code of *tnbF*'s A domains and predict specificity. The web-based software NRPS-PKS (Ansari et al., 2004) is a knowledge-based tool based on a comprehensive analysis of sequence and structural features of experimentally characterized biosynthetic gene clusters. It uses a BLAST analysis to identify domains and uses an alignment with structural templates to identify the specificity-conferring code of a given A domain., An updated version of the NRPS-PKS release from Ansari et al. (2004) is hosted on the SBSPKS server (Anand et al., 2010). The NRPSpredictor2 (Rottig et al., 2011) is an improved version of the program originally released by Rasuch et al. (2005). This program uses a machine learning method, transductive support vector machines (TSVM), based on the physico-chemical characteristics of the substrate binding pocket to identify an A domain's specificity-conferring code. The third online tool is the PKS/NRPS analysis website described by Bachmann and Ravel (2009). The PKS/NRPS analysis website relies on Hidden Markov Model (HMM) searches, which are statistical representations of a group of proteins using a position-specific scoring system.

The three online tools predicted the same specificity for the A domain of the second module (M2) of *tnbF*: a specificity-conferring code of DVWHFSLV for the activation of serine (Table 2.4). However, the prediction for the A domain of first

module (M1) was inconclusive. The NRPS-PKS tool and PKS/NRPS analysis website reported a similar specificity-conferring code for this A domain, while the NRPSpredictor2 reported a different code (Table 2.4). More importantly, though, all three programs failed to identify a reliably significant prediction for this A domain.

Table 2.4. Specificity codes of *tnbF*'s A domains predicted by three separate online tools. (-) indicates an undetermined residue at that particular position.

	Prediction tool		
	NRPS-PKS	**PKS/NRPS analysis website**	**NRPSpredictor2**
M1 A domain	D-WDIILV	D-WDIILV	DSDDGGLV
M2 A domain	DVWHFSLV	DVWHFSLV	DVWHFSLV

To further analyze the specificity prediction of *tnbF*'s A domains, a phylogenetic analysis of the specificity-conferring codes of *tnbF*'s A domains was carried out with a selection of A domains with annotated specificity. The work by Stachelhaus et al. (1999) and Challis et al. (2000) shows that phylogenetic analysis of the specificity-conferring codes from A domains reveals that A domains group according to substrate specificity. As reasoned by Challis et al. (2000), an eight amino acid specificity code was used in the current analysis as opposed to the 10 amino acid code suggested by Stachlehaus et al. (1999). Briefly, Cys331 of GrsA points away from the binding pocket and Lys517 is strictly conserved among A domains. The results of the present analysis show that the M2 A domain groups with other serine-activating A domains, as expected (Figure 2.2). The M1 A domain groups with lysine, a hydrophilic amino acid with a long, positively charged side chain.

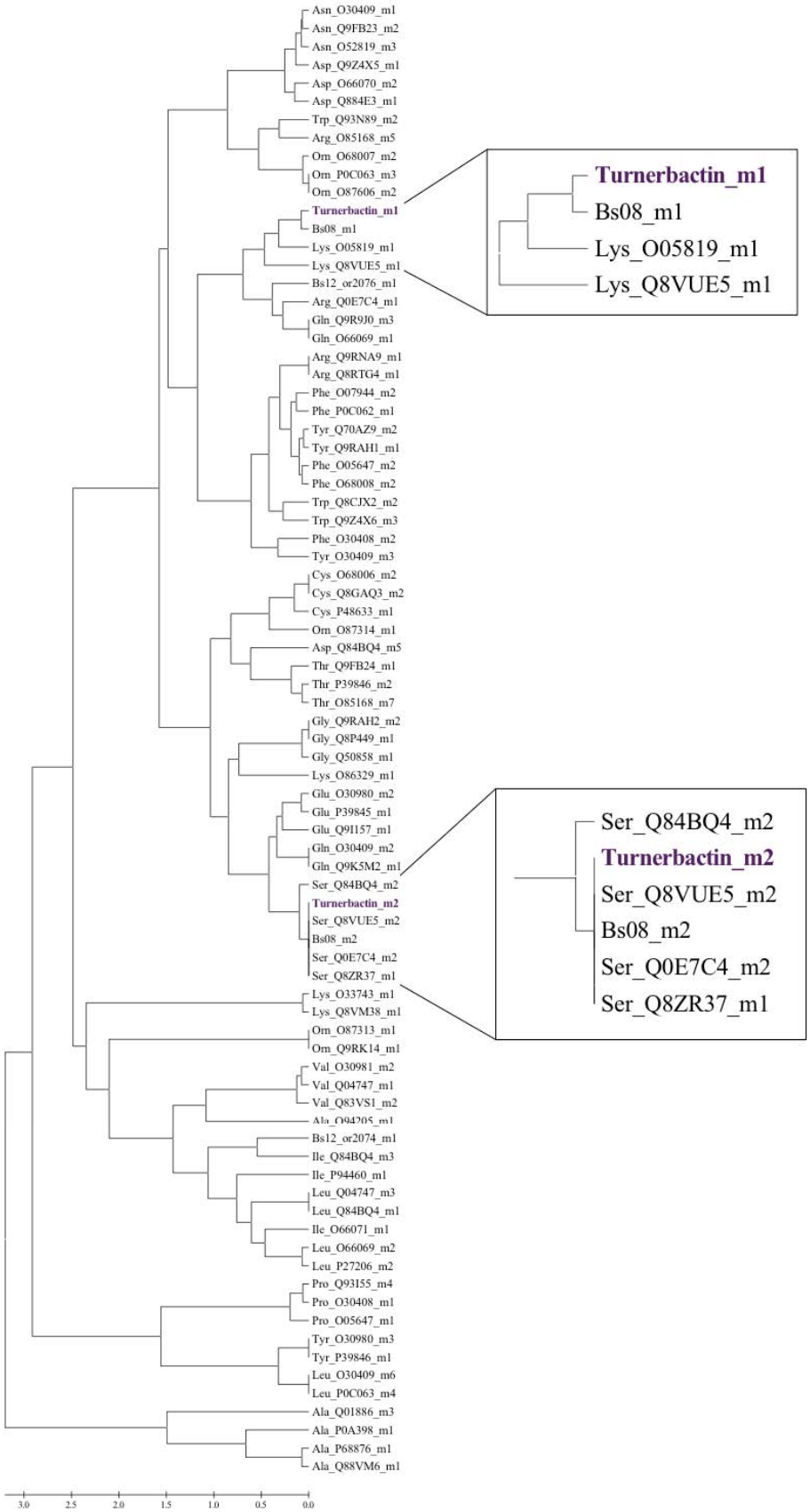
Asn_O30409_m1
Asn_Q9FB23_m2
Asn_O52819_m3
Asp_Q9Z4X5_m1
Asp_O66070_m2
Asp_Q884E3_m1
Trp_Q93N89_m2
Arg_O85168_m5
Orn_O68007_m2
Orn_P0C063_m3
Orn_O87606_m2
Turnerbactin_m1
Bs08_m1
Lys_O05819_m1
Lys_Q8VUE5_m1
Bs12_or2076_m1
Arg_Q0E7C4_m1
Gln_Q9R9J0_m3
Gln_O66069_m1
Arg_Q9RNA9_m1
Arg_Q8RTG4_m1
Phe_O07944_m2
Phe_P0C062_m1
Tyr_Q70AZ9_m2
Tyr_Q9RAH1_m1
Phe_O05647_m2
Phe_O68008_m2
Trp_Q8CJX2_m2
Trp_Q9Z4X6_m3
Phe_O30408_m2
Tyr_O30409_m3
Cys_O68006_m2
Cys_Q8GAQ3_m2
Cys_P48633_m1
Orn_O87314_m1
Asp_Q84BQ4_m5
Thr_Q9FB24_m1
Thr_P39846_m2
Thr_O85168_m7
Gly_Q9RAH2_m2
Gly_Q8P449_m1
Gly_Q50858_m1
Lys_O86329_m1
Glu_O30980_m2
Glu_P39845_m1
Glu_Q9I157_m1
Gln_O30409_m2
Gln_Q9K5M2_m1
Ser_Q84BQ4_m2
Turnerbactin_m2
Ser_Q8VUE5_m2
Bs08_m2
Ser_Q0E7C4_m2
Ser_Q8ZR37_m1
Lys_O33743_m1
Lys_Q8VM38_m1
Orn_O87313_m1
Orn_Q9RK14_m1
Val_O30981_m2
Val_Q04747_m1
Val_Q83VS1_m2
Bs12_or2074_m1
Ile_Q84BQ4_m3
Ile_P94460_m1
Leu_Q04747_m3
Leu_Q84BQ4_m1
Ile_O66071_m1
Leu_O66069_m2
Leu_P27206_m2
Pro_Q93I55_m4
Pro_O30408_m1
Pro_O05647_m1
Tyr_O30980_m3
Tyr_P39846_m1
Leu_O30409_m6
Leu_P0C063_m4
Ala_Q01886_m3
Ala_P0A398_m1
Ala_P68876_m1
Ala_Q88VM6_m1
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Turnerbactin_m1
Bs08_m1
Lys_O05819_m1
Lys_Q8VUE5_m1
Ser_Q84BQ4_m2
Turnerbactin_m2
Ser_Q8VUE5_m2
Bs08_m2
Ser_Q0E7C4_m2
Ser_Q8ZR37_m1

Figure 2.2. Phylogenetic clustering of A domains using the unweighted pair group method using arithmetic averages (UPGMA). The A domains from *tnbF* are shown in purple. Each A domain is labeled with its amino acid specificity, followed by its UniProtKB accession number, followed by the module number of the NRPS from which the A domain is referring.

2.3.2.2 C domain. Rausch et al. (2007) demonstrated that the reconstructed phylogeny of condensation domains shows a functional grouping of C domains, rather than a species phylogeny or substrate specificity. A similar phylogenetic analysis was carried out with the C domains from *tnbF*.

The results of the maximum likelihood analysis show an overall similar phylogeny to that of Rausch et al. (2007), with Dual E/C and $^{D}C_{L}$ domains grouping closely together as well as Starter and $^{L}C_{L}$ domains grouping together (Figure 2.3). The M1 C domain from *tnbF* groups with other Starter domains, as expected from its location in the NRPS. The M2 C domain of *tnbF* groups within the $^{D}C_{L}$ functional domains and this placement is strongly supported, evidenced by the high bootstrap values at the base of the Dual E/C and $^{D}C_{L}$ groups. In most cases, $^{D}C_{L}$ domains are preceded by an E domain. However, no E domain was detected in *tnbF* through pHMM analysis and no external racemase was detected in the gene cluster. This suggests that the M2 C domain of *tnbF* has $^{L}C_{L}$ activity as opposed to its predicted $^{D}C_{L}$ activity.

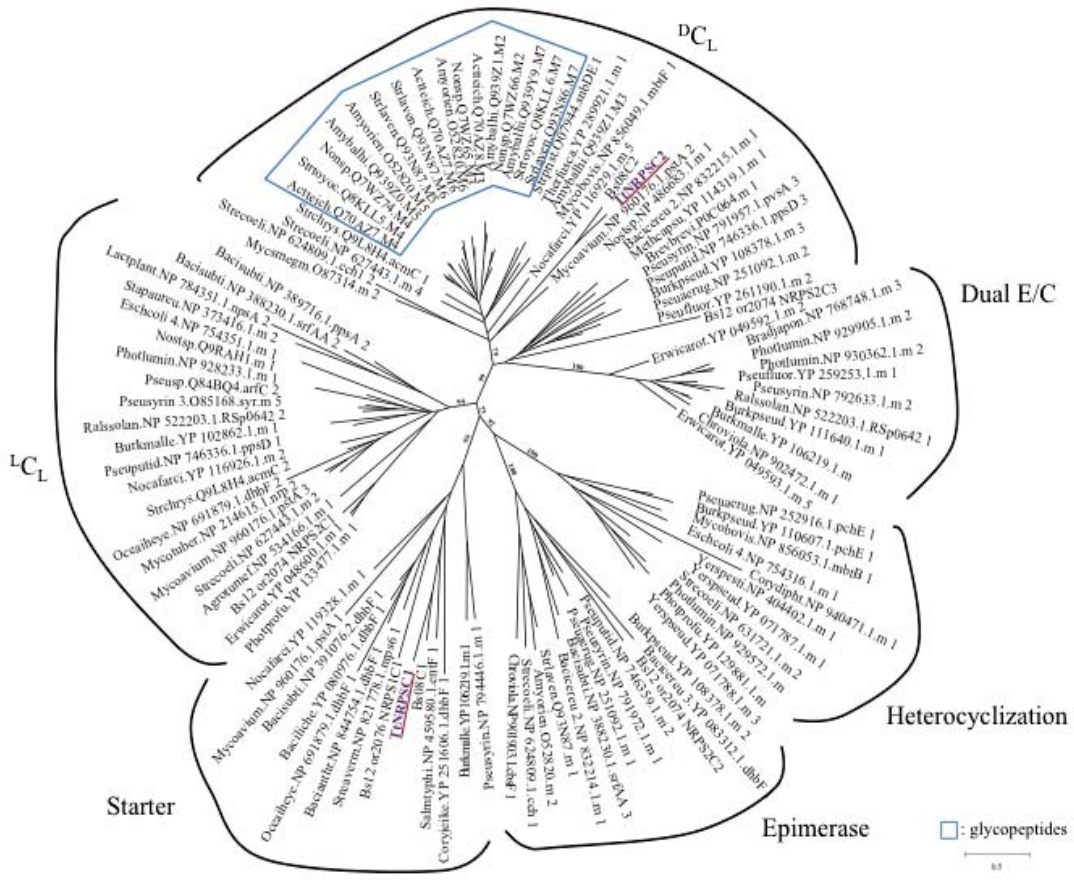

Figure 2.3. Maximum likelihood tree of C domains, showing the grouping of different C domain subtypes. The C domains of *tnbF* are in purple and underlined in red. *tnbF*'s M2 C domain groups with the $^{D}C_{L}$ functional group, while *tnbF*'s M2 C domain groups with the Starter functional group. Each C domain is labeled with the organism name, followed by accession number, followed by the module number from which the C domain is referring. Bootstrap values are based on 100 replicates and are only shown for the basal branches of groups.

2.3.2.3 TE domain. The crystal structure of two NRPS TE domains have been reported: the surfactin TE (SrfTE) (Bruner et al., 2002) and the fengycin TE (FenTE) (Samel et al., 2006). Primary sequence data has been compared with structural data to reveal locations and potential functions of key residues. One particular residue, a Pro located near the oxyanion hole in the active site of the TE domain, Pro26 in SrfTE, has been shown to influence peptide release (Tseng et al., 2002). When compared with lipases, a family of ester/lactone hydrolyzing enzymes, the sequence alignment

showed that this Pro residue was conserved among TE domains, while conserved as a Gly in lipases. Mutation of this residue to a Gly in SrfTE led to a significant change in product ratio, favoring hydrolysis over cyclization (Tseng et al., 2002). This suggests that the Pro residue imposes a conformational constraint on the TE domain that favors an active site protected from water, allowing for intramolecular cyclization. The flexible Gly may allow water access to the active site and lead to increased hydrolysis. The TE domain of *tnbF* was analyzed for the identity of this key residue.

A multiple sequence alignment was constructed with the TE domain from *tnbF* and other annotated TE domains. The results show that the *tnbF* TE domain contains a Pro residue in the conserved location proposed to influence peptide release from the NRPS (Figure 2.4). This suggests that *tnbF* produces a cyclic peptide product.

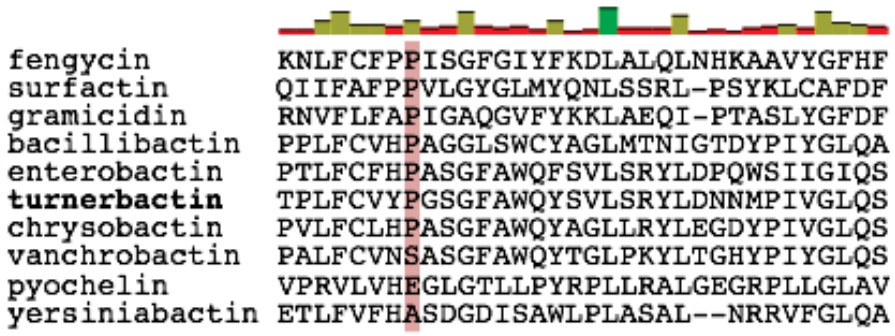

Figure 2.4. Partial alignment of TE domains. TE domain sequences were extracted from annotated NRPS's. UniProtKB accession numbers of NRPS's are as follows: fengycin, L42523; surfactin, Q08787; gramicidin, P0C064; bacillibactin, P45745; enterobactin, P11454; chrysobactin, Q8VUE5; vanchrobactin, Q0E7C4; pyochelin, O85740; yersiniabactin, Q9Z373. The residue highlighted in red represents the conserved Pro26 in SrfTE, hypothesized to indicate an intramolecular cyclization mechanism of release by TE domains.

2.3.3 Chemical characterization of the siderophore

The structures of the siderophores isolated from *T. turnerae* T7901 are shown in Figure 2.5. The siderophores from *T. turnerae* were purified from iron-deficient liquid cultures of shipworm basal medium (SBM) by enriching siderophores from culture supernatants on an HP20 column, followed by purification by RP-HPLC. The CAS and Arnow assay were used to track the siderophores throughout the purification process. RP-HPLC revealed five peaks displaying CAS and Arnow activity (Figure 2.6). High resolution electrospray ionization mass spectrometry (HRESIMS) determined the mass of the molecular ion $[M+H]^+$: turnerbactin, m/z 356.1454, corresponding to a molecular formula of $C_{15}H_{22}N_3O_7$ (calculated 356.1458); diturnerbactin, m/z 693.2736, corresponding to a molecular formula of $C_{30}H_{41}N_6O_{13}$ (calculated 693.6723); dehydrated diturnerbactin, m/z 675.2626, corresponding to a molecular formula of $C_{30}H_{39}N_6O_{12}$ (calculated 675.2627); triturnerbactin, m/z 1030.4003, corresponding to a molecular formula of $C_{45}H_{60}N_9O_{19}$ (calculated 1030.4007); dehydrated triturnerbactin, m/z 1012.3905, corresponding to a molecular formula of $C_{45}H_{58}N_9O_{18}$ (calculated 1012.3902).

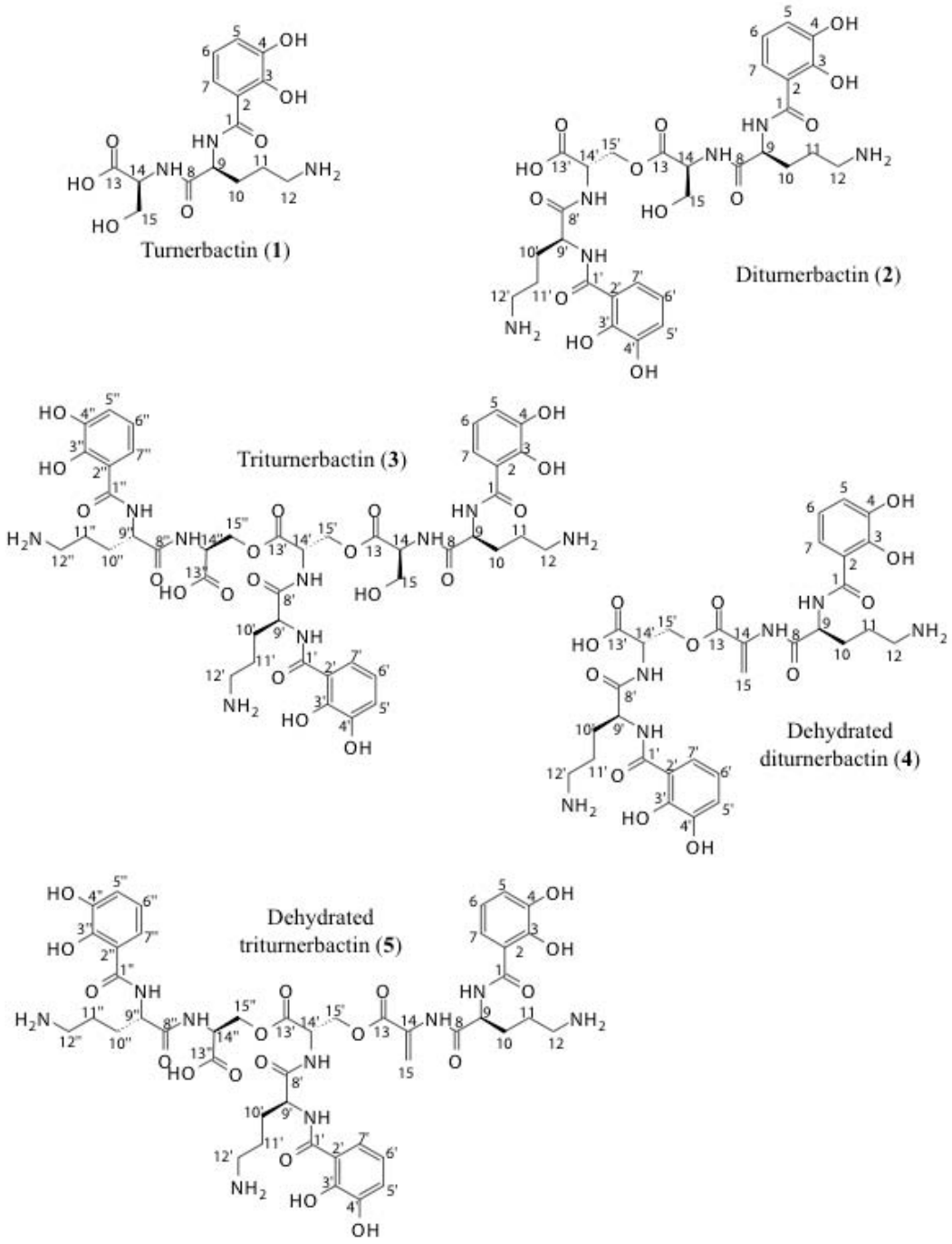

Figure 2.5. Structures of siderophores isolated from *T. turnerae* T7901.

ESIMS/MS analysis of each of these compounds is summarized in Table 2.5. All compounds exhibited similar fragmentation patterns, with overlap of fragments

when applicable. The masses of the fragments could be correlated with the loss of various constituents of the siderophore, indicated in Table 2.5.

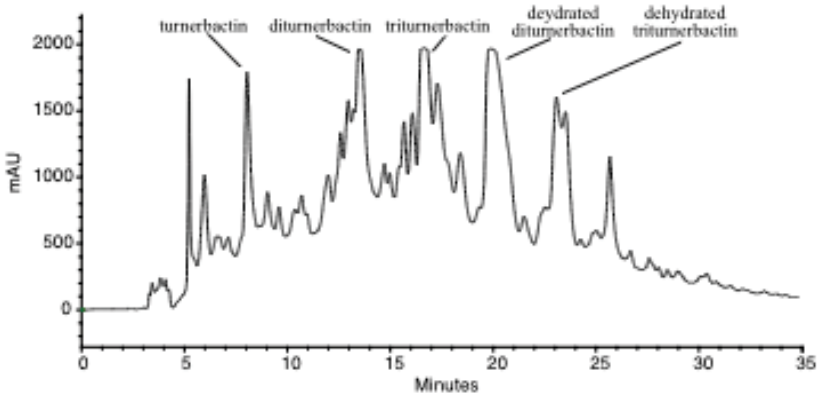

Figure 2.6. HPLC trace of HP20 extract of *T. turnerae* T7901 culture supernatant, recorded at 215 nm.

Table 2.5. Molecular ions and common mass fragments of siderophore from *T. turnerae* T7901. Fragment losses refer to the compound listed immediately above in the table.

Dehydrated triturnerbactin (5)	**Dehydrated diturnerbactin (4)**	**Triturnerbactin (3)**	**Diturnerbactin (2)**	**Turnerbactin (1)**	**Fragment**
1012.4	675.2	1030.4	693.3	356.1	Parent ion
762.3		780.3			Loss of DHB-Orn
675.2		693.3	675.2		Loss of Ser
425.2	425.2	443.2	443.2		Loss of DHB-Orn
338.1	338.1	356.1	356.1		Loss of Ser
251.1	251.1	251.1	251.1	251.1	DHB-Orn
115.1	115.1	115.1	115.1	115.1	Orn

The 1-D ^{1}H and ^{13}C NMR assignments of **1** were confirmed by a variety of 2-D experiments ^{1}H-^{1}H TOCSY, HSQC, and HMBC (Table 2.6). The ^{13}C NMR spectrum shows 15 distinct C resonances corresponding to three carbonyl carbons (δ 170.0 to 172.1), four methylene carbons (δ 23.4 to 38.9 for ornithine, δ 61.4 for

serine), two methine carbons (δ 52.3 for ornithine, δ54.8 for serine), and six aromatic carbons (δ 116.0 to 147.8). The ^{1}H NMR spectrum shows 12 distinct resonances corresponding to three aromatic protons (δ 6.77 to 7.36), seven methylene protons (δ 1.79 to 3.00 ornithine, δ 3.88, 3.98 for serine), and two methine protons (δ 4.82 for ornithine, δ 4.55 for serine). The aromatic splitting pattern in the ^{1}H NMR spectrum is indicative of 2, 3-DHB moiety. A HMBC correlation between the α-proton of ornithine and the carbonyl carbon of DHB indicates that DHB is attached to the α-amine group of ornithine. A HMBC correlation between the α-proton of serine and the carbonyl carbon of ornithine confirms the ornithine-serine peptide bond.

The structures of **2** and **3** were inferred using MS, ^{1}H NMR, and 2-D NMR experiments HSQC and HMBC. The NMR assignments for 3 are shown in Table 2.5, the NMR assignments for 2 are shown in Appendix A. The ^{1}H NMR spectra of **2** and **3** are similar to that of **1**, with the addition of a downfield shift of serine methylene protons of **2** and **3** (δ 4.43 to 4.75) compared to the serine methylene protons of **1** (δ 3.88, 3.98), indicating the presence of serine ester linkages in **2** and **3**.

The MS parent ion masses **5** suggested a cyclic trimer as the result of an additional serine ester bond, leading to a triserine lactone backbone. However, ^{1}H and ^{13}C NMR in addition to 2-D experiments ^{1}H-^{1}H TOCSY, HSQC, and HMBC revealed a dehydrated analogue of the linear trimer (Table 2.5). The dehydration of the serine hydroxyl group resulted in an alkene, corresponding to a carbon resonance of δ 112.81 and proton resonances of δ 6.12 and 5.80. Similar spectra were obtained for **4**. The NMR assignments for 4 are shown in Appendix A.

All NMR spectra for compounds **1-5** are shown in Appendix A.

Table 2.6. NMR data for Turnerbactin, Triturnerbactin, and Dehydrated triturnerbactin (800 MHz) in CD_3OD.

	Turnerbactin (1)				Triturnerbactin (3)		
Position	**δ_C**	**δ_H (*J* in Hz)**	**TOCSY**	**HMBC**	**δ_H (*J* in Hz)**	**HSQC**	**HMBC**
DHB							
1, 1', 1"	169.98						
2, 2', 2"	116.01						
3, 3', 3"	147.81						
4, 4', 4"	145.82						
5, 5', 5"	118.38	6.98, dd (1.6, 8.0), [1H]	6, 7	2, 3, 4, 6, 7	6.97, m, [3H]	118.5	2, 3, 4, 6, 7
6, 6', 6"	118.52	6.77, t (8.0), [1H]	5, 7	1, 2, 3, 4, 5, 7	6.76, m, [3H]	118.7	1, 2, 3, 4, 5, 7
7, 7', 7"	118.44	7.36, dd (0.8, 8.0), [1H]	5, 6	1, 2, 3, 4, 5, 6	7.36, m, [3H]	118.6	1, 2, 3, 4, 5, 6
Ornithine							
8, 8', 8"	172.12						
9	52.26	4.82, dd (5.6, 8.0), [1H]	10, 11, 12	1, 8, 10, 11	4.71, m, [1H]	52.6	1, 8, 10, 11
9'					4.77, m, [1H]	52.5	1', 8', 10', 11'
9"					4.79, m, [1H]	52.2	1", 8", 10", 11"
10, 10', 10"	29.06	2.09, m [1H] 1.89, m [1H]	9, 10, 11, 12 9, 10, 11, 12	8, 9, 11, 12	2.05, m, [3H] 1.89, m, [3H]	28.7	8, 9, 11, 12 8, 9, 11, 12
11, 11', 11"	23.39	1.83, m, [1H] 1.79, m [1H]	9, 10, 11, 12 9, 10, 11, 12	9, 10, 12	1.81, m, [6H]	28.14, 28.13	9, 10, 12
12, 12', 12"	38.93	3.00, m, [2H]	9, 10, 11, 12	10, 11	2.99, m, [6H]	38.9	10, 11
Serine							
13	171.91						
13'							
13"							
14	54.82	4.55, t (8.0,) [1H]	15	13, 15	4.56, m, [1H]	55.0	8, 13, 15
14'					4.80, m, [1H]	52.0	8', 13', 15'
14"					4.87, m, [1H]	51.7	8", 13", 15"
15	61.35	3.98, dd (4.8, 11.2), [1H] 3.88, dd (4.0, 11.2), [1H]	14, 15	13, 14	3.95, m, [1H] 3.82, m, [1H]	62.2	13, 14 13, 14
15'					4.72, m, [1H] 4.43, m, [1H]	64.5	13, 13', 14' 13, 13', 14'
15"					4.57, m, [1H] 4.49, m, [1H]	63.5	13", 14" 13", 14"

Dehydrated triturnerbactin				
Position	**δ_C**	**δ_H (*J* in Hz)**	**TOCSY**	**HMBC**
DHB				
1, 1', 1"	170.74 170.88 170.98			
2, 2', 2"	117.21 117.16 117.01			
3, 3', 3"	149.33 149.36 149.50			
4, 4', 4"	147.24 147.15			
5, 5', 5"	119.97 119.91 119.90	6.95, dd (1.6, 8.0), [1H] 6.93, dd (0.8, 8.0), [1H] 6.99, dd (1.6, 8.0), [1H]	6, 7	2, 3, 4, 6 2', 3', 4', 6' 2", 3", 4", 6"
6, 6', 6"	120.04 120.03	6.72, q (8.0), [2H] 6.77, t (8.0), [1H]	5, 7	2, 3, 4, 5, 7
7, 7', 7"	119.81 119.76 119.69	7.33, dd (0.8, 8.0), [1H] 7.30, dd (0.8, 8.0), [1H] 7.36, dd (0.8, 8.0), [1H]	5, 6	1, 3, 4, 6 1', 3', 4', 6' 1", 3", 4", 6"
Ornithine				
8, 8', 8"	173.66, 173.89, 172.23			
9	53.95	4.75, m, [1H]	10, 11, 12	1, 8, 10, 11
9'	54.09	4.67, m, [1H]		1', 8', 10', 11'
9"	54.41	4.84, m, [1H]		1", 8", 10", 11"
10, 10', 10"	29.99, 29.79, 29.70	2.05, m, [3H] 1.89, m, [3H]	9, 11, 12	8, 9, 11, 12 8, 9, 11, 12
11, 11', 11"	25.09, 25.01, 25.00	1.83, m, [3H] 1.79, m, [3H]	9, 10, 12	9, 10, 12 9, 10, 12
12, 12', 12"	40.31	3.00, m, [6H]	9, 10, 11	10, 11
Serine				
13	164.4			
13'	169.91			
13"	172.29			
14	132.7			
14'	52.99	4.89, m, [1H]	15'	8', 13', 15'
14"	53.22	4.82, m, [1H]	15"	8", 13", 15"
15	112.81	6.12, s, [1H] 5.80, s, [1H]	15 15	13, 14 13, 14
15'	65.32	4.56, m, [1H] 4.53, m, [1H]	15', 14' 15', 14'	13, 13', 14' 13, 13', 14'
15"	65.88	4.74, m, [1H] 4.38, dd (6.4, 11.2), [1H]	15", 14" 15", 14"	13', 13", 14" 13', 13", 14"

Amino acid analysis by Marfey's method established the presence of L-ornithine and L-serine in **3**. The HPLC trace of the hydrolysate of triturnerbactin shows peaks with the same retention time as L-ornithine and L-serine. These peaks were checked by MS to confirm their identity.

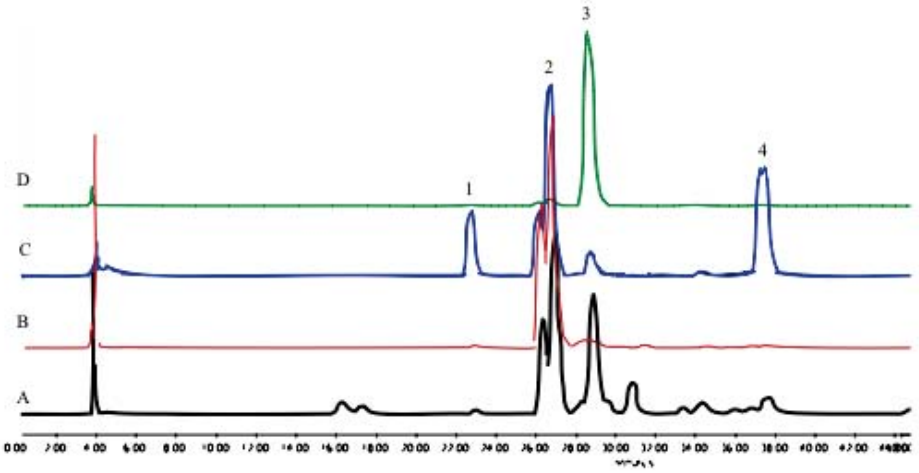

Figure 2.7. HPLC chromatograms of FDAA amino acid derivatives. (A) Trace of the hydrolysate of triturnerbactin in black, (B) trace of L-ornithine in red, (C) trace of DL-ornithine in blue, (D) trace of L-serine in green. (1) elution of D-ornithine, (2) elution of L-ornithine, (3) elution of L-serine, (4) elution of Marfey's reagent (FDAA).

2.3.4 Disruption of *tnbF* diminishes siderophore activity

In order to correlate the *tnb* gene cluster with the biosynthesis of turnerbactin, a *tnbF* mutant was constructed by integrating a chloramphenicol (Cm) resistance cassette into the M1 C domain of *tnbF* by a single crossover recombination (Figure 2.1). The resulting strain was named TtAH03.

Wild-type *T. turnerae* T7901 grown in iron-deficient SBM medium produces a yellow pigmentation of the medium, presumably the result of the excreted

siderophore in culture supernatant as this color can be tracked with the siderophores during the purification process. TtAH03 grown in iron-deficient SBM medium does not produce this yellow color (Figure 2.8). Growth of TtAH03 in iron-deficient SBM medium was inhibited (Figure 2.9) and siderophore production (measured by the CAS assay) was severely reduced (Figure 2.10). The residual CAS activity shown for TtAH03 is presumably due to the DHB units produced by the genes *tnbCEBA*. These results support the hypothesized role of the *tnb* gene cluster in the role of turnerbactin biosynthesis.

Figure 2.8. Iron-deficient cultures of T7901 (left) and TtAH02(right).

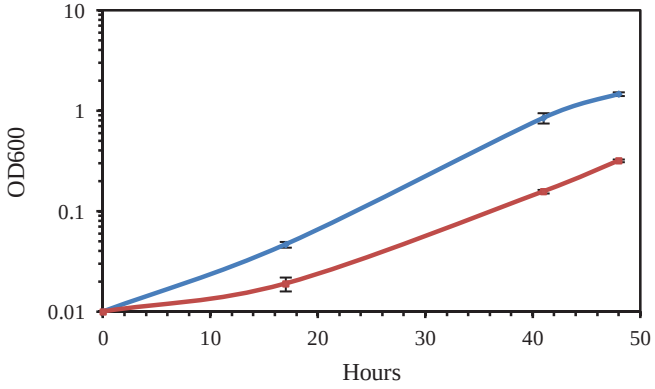

Figure 2.9. Growth of T7901 (blue) and TtAH03 (red) in iron-deficient SBM medium.

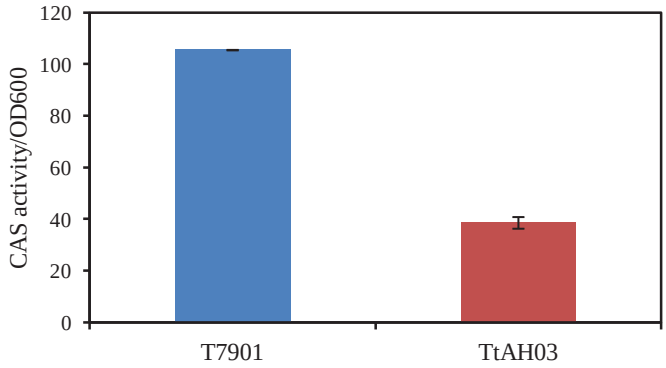

Figure 2.10. Siderophore production of wild-type T7901 compared to *tnbF* mutant TtAH03. Disruption of *tnbF* leads to a significant decrease in siderophore activity. Siderophore production was measured with the CAS assay and presented is CAS activity normalized to OD600 measurements of each culture.

2.3.5 Detection of turnerbactin in shipworm samples

The presence of the turnerbactin biosynthetic gene cluster in the genome of *T. turnerae* T7901 does not implicate its utility in symbiosis. Detection of the compound in the shipworm host would provide evidence that the compound is produced and utilized in the symbiosis. The presence of turnerbactin in the shipworm host was investigated using MS. The shipworm, *Lyrodus pedicellatus*, is known to contain the symbiont *T. turnerae* (Distel et al., 1991; Distel et al., 2002a). A crude, methanolic extract of whole *L. pedicellatus* tissue samples was analyzed by LC-HRMS/MS and compared to a pure turnerbactin standard. The method used in this analysis shows pure turnerbactin eluting from the column over the span of approximately 7.5 – 8.5 min. The HRESIMS m/z peak for this elution product was 356.1446. Tandem MS analysis shows the presence of two daughter ions with m/z values of 338.2 and 251.1. Analysis of the *L. pedicellatus* extract shows a peak eluting at 7.94 min with a m/z value of 356.1447. Tandem MS analysis of this peak shows the presence of the two daughter ions at 338.2 and 251.1. These results

provide strong evidence for the presence of turnerbactin in these samples of *L. pedicellatus* and suggest that turnerbactin is produced in the symbiosis.

2.4 Discussion

This work presents the purification, structural characterization, and biosynthesis genes for a novel siderophore, turnerbactin, from the shipworm symbiont *T. turnerae* T7901. Turnerbactin was isolated as a monomer, dimer, dehydrated dimer, linear trimer, and dehydrated trimer. The linear trimer, triturnerbactin, is most likely produced by the NRPS encoded by *tnbF* in such a way that the modules are used three consecutive times to assemble a trimer of 2,3-DHB-L-ornithyl-L-serine on the M2 PCP domain, which is then hydrolyzed by the terminal TE domain to release the linear peptide product. The proposed biosynthesis is shown in Figure 2.11. DHB is activated by EntE and then transferred to the aryl carrier C-terminal portion of EntB. L-ornithine is activated by the M1 A domain of TnbF and is condensed to DHB by the M1 C domain, forming a DHB-Orn intermediate on the M1 PCP domain. L-serine is activated by the M2 A domain and the M2 C domain condenses the Ser to the DHB-Orn intermediate to form a monomer unit of DHB-Orn-Ser-S-PCP thioester intermediate on the M2 PCP domain. A conserved Ser residue in the TE domain mounts a nucleophilic attack of the PCP domain-bound thioester intermediate, leading to a DHB-Orn-Ser-O-TE ester intermediate. After another round of DHB-Orn-Ser biosynthesis, the serine side chain hydroxyl group of the DHB-Orn-Ser-O-TE intermediate attacks the DHB-Orn-Ser-S-PCP to yield a $(DHB\text{-}Orn\text{-}Ser)_2$-S-PCP, which is then transferred to the TE domain. A third consecutive round yields a

$(DHB\text{-}Orn\text{-}Ser)_3$-O-TE intermediate, which is then hydrolyzed by water to yield the linear trimer $(DHB\text{-}Orn\text{-}Ser)_3$. It is uncertain whether the other compounds discovered in this study are byproducts from regulated biosynthesis by the NRPS, incomplete biosynthesis by the NRPS, enzymatic degradation by an esterase, or hydrolysis products from either the purification process or culture conditions.

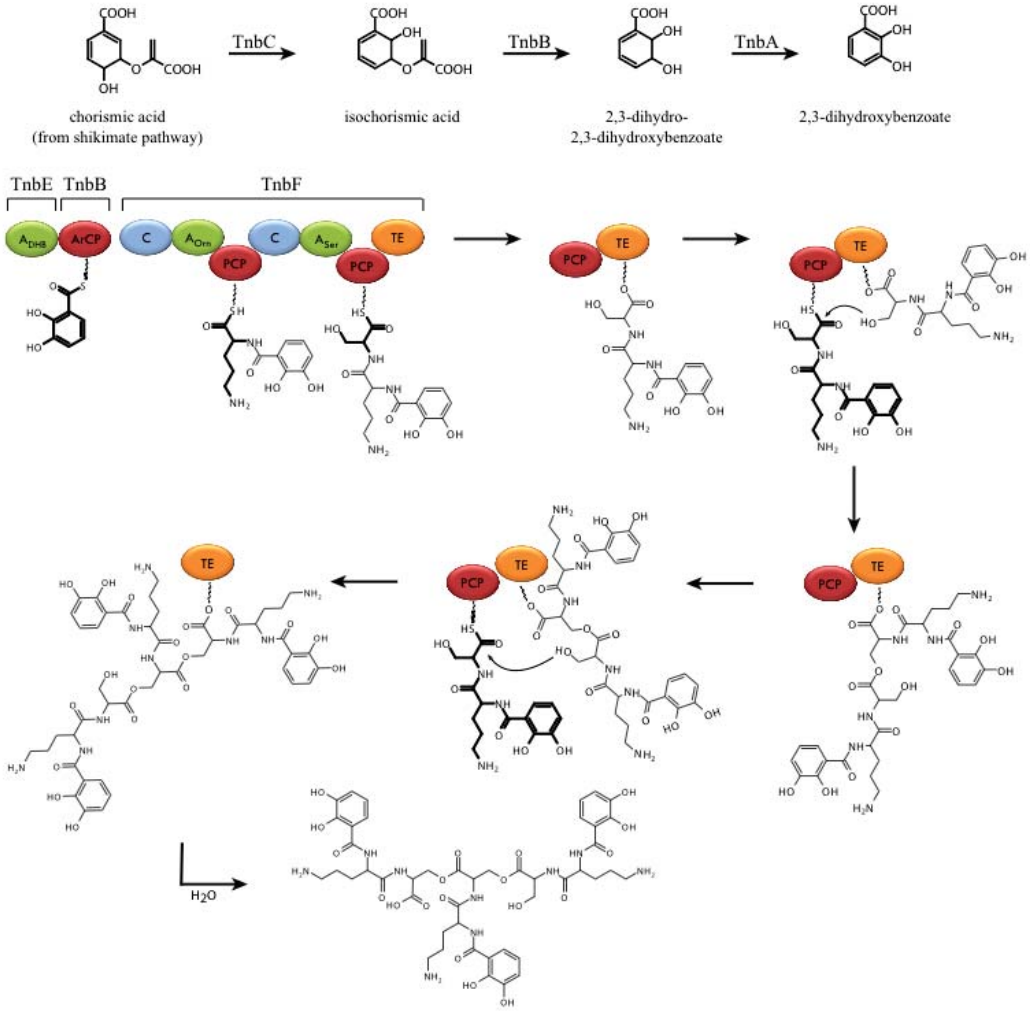

Figure 2.11. Proposed biosynthesis of triturnerbactin by TnbCEBAF.

The cyclic trimer of turnerbactin was not found in this study. The cyclic trimer may be produced by *T. turnerae*, as predicted by the TE domain analysis suggesting the production of a cyclic compound, but not detected under these conditions due to instability or degradation. On the other hand, the bioinformatic prediction may be erroneous. The bioinformatic analyses of *tnbF*'s catalytic domains suggest these domains elude accurate prediction. Predictive analysis of A domains by three separate programs fail to suggest a specificity with reasonable cut-off scores. Phylogenetic analysis groups the M2 A domain of *tnbF* with a substrate with similar properties, Lys, but this A domain did not group with other Orn-activating A domains even though the chemical analysis clearly shows Orn is present. C domain analysis suggests a $^{D}C_{L}$ acting domain, while chemical data suggest a $^{L}C_{L}$ acting domain. Thus, bioinformatic analysis of the TE domain may not accurately reflect the true activity of the *tnbF* TE domain either. Taken all together, the addition of *tnbF* to the databases of NRPS domains may help to improve the functional prediction of as-yet-undiscovered NRPS domains.

The NRPS's responsible for the production glycopeptides, found in actinomycetes, also exhibit aberrant C domain prediction. All glycopeptides NRPS's contain seven modules. Rausch et al. (2007) showed that the 4th module (M4) and M7 C domains act as $^{L}C_{L}$ domains while clustering in the $^{D}C_{L}$ group of C domains in phylogenetic analysis. While the M2 C domain of *tnbF* does not cluster directly with these glycopeptides C domains (Figure 2.3), a similar change of function is assumed to have occurred, most likely a result of convergent evolution. Occasionally, an external racemase can be found in a biosynthetic gene cluster, providing a D-amino

acid for the NRPS. This is the case for cyclosporine, where an external racemase provides D-Ala for the first module of the cyclosporine synthetase (Hoffmann et al. 1994). However, no external racemase was detected in the turnerbactin biosynthetic gene cluster and no D-Orn was detected in amino acid analysis. It is possible that an epimerase domain was once a part of the tunerbactin NRPS but was lost at some point in time.

Nevertheless, turnerbactin represents a novel siderophore, structurally similar to the catecholate siderophores vanchrobactin and chrysobactin. A catecholate siderophore was partially characterized from the soil diazotroph, *Azospirillum brasilense*, which contained equimolar amounts of 2,3-DHB, ornithine, and serine (Bachhawat and Ghosh, 1987). However, a structure was never elucidated or presented. Vanchrobactin, 2,3-DHB-D-arginyl-L-serine, was originally described from the fish pathogen *Vibrio anguillarum* serotype O2 (Soengas et al., 2006). Chrysobactin, 2,3-DHB-D-lysyl-L-serine, was originally described from the plant pathogen *Erwinia chrysanthemi* 3937 (Persmark et al., 1989). Along with turnerbactin, these three siderophores share a 2,3-DHB functional moiety, a hydrophilic long, positively charged spacer amino acid, and a serine backbone. However, the biosynthesis of turnerbactin differs from vanchrobactin and chrysobactin in that the spacer amino acid is not epimerized to a D configuration, due to the absence of an E domain in *tnbF* (Figure 2.12). The similarity of NRPS modules and the finding of a $^{D}C_{L}$-like C domain for the M2 of *tnbF* suggest that an E domain may have been lost at some point.

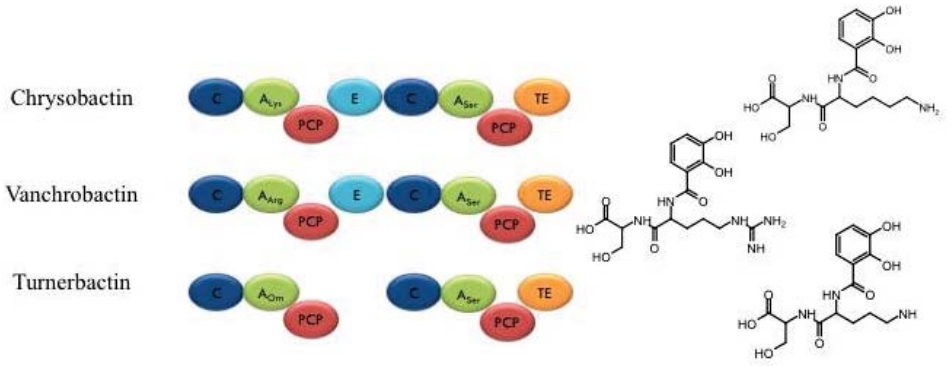

Figure 2.12. NRPS organization for chrysobactin, vanchrobactin, and turnerbactin and their monomer unit products. C: condensation domain, A: adenylation domain, PCP: peptidyl carrier protein, E: epimerase domain, TE: thioesterase domain.

Until recently, only the monomer units of vanchrobatin and chrysobactin had been reported. Sandy et al. discovered the trimeric form of vanchrobactin from an open ocean bacterial isolate *Vibrio* sp. DS40M4 (Sandy et al., 2010), along with the monomeric and dimeric vanchrobactin. The cyclic trimer was not found, and bioinformatic analysis of the TE domain involved in vanchrobactin biosynthesis in *Vibrio* sp. DS40M4 supports the release of a linear trimer (Appendix B). Soon thereafter, Sandy and Butler (2011) reported the cyclic trimer of chrysobactin from the plant pathogen *Dickeya chrysanthemi* EC16. These findings beg the question: what determines whether an NRPS will act linearly, using its modules once to produce a peptide product, or iteratively, using its modules in multiple subsequent rounds to produce a single oligomeric peptide product. It has been proposed that the C-terminal TE domain is the key to determining a linear versus iterative process, but multiple sequence alignments (Mootz et al., 2002), structural analyses (Samel et al., 2006), and in vitro experiments (Hoyer et al., 2007) have failed to explain the mechanism behind this process.

When comparing a cyclic compound to its linear counterpart, a cyclic compound will likely have a higher stability constant, as the flexibility of the cyclic ligand and its corresponding iron complex will be less than that of the linear ligand, thereby decreasing the entropy difference. This is seen in the stability constants for cyclic enterobactin and its linearized form, the linear trimer of DHB-serine (Scarrow et al., 1991). The calculated stability constant of linear enterobactin is 10^{43}, compared to the calculated stability constant of 10^{49} for cyclic enterobactin. However, the linear trimer still provides a siderophore with relatively strong affinity for iron and the initial rates of uptake for both the linear trimer and linear dimer ferric enterobactin complexes are essentially the same as the cyclic ferric enterobactin complex (Scarrow et al., 1991).

The triscatecholate class of siderophores contains the highest stability constants measured for siderophores to date, including enterobactin, 2,3-DHB-Ser (10^{49}) (Loomis and Raymond, 1991) and bacillibactin, 2,3-DHB-Gly-Thr ($10^{47.6}$) (Dertz et al., 2006). Both siderophores are three-fold symmetrical, hexadentate catecholate ligands, but the major difference is the presence of a spacer unit connecting the catecholate functional groups to the trilactone backbone in bacillibactin. The spacer unit is slightly detrimental to the overall stability of the ferric complex, as suggested by the lowered thermodynamic stability for bacillibactin (Dertz et al., 2006). Therefore, the ability of the linear trimer of turnerbactin to bind ferric iron may be slightly diminished when compared to enterobactin, due to the presence of a spacer amino acid, and bacillibactin, due to its linear form. Thermodynamic stability measurements of the ferric turnerbactin complex would

more accurately allow for comparison of ferric binding abilities amongst these triscatecholate siderophores.

A compound was detected in *L. pedicellatus* extracts that matches the MS and tandem MS profile of turnerbactin. The compound detected in the shipworm shared the same LC retention time, high resolution molecular ion mass, and daughter ions in tandem MS fragmentation as pure turnerbactin. The detection of turnerbactin in a shipworm sample suggests that the siderophore is utilized in some capacity in the symbiosis. Since whole animal tissue was used in these experiments, it is not known whether these siderophores are confined to the immediate vicinity of the symbionts in the gills, or if they appear in other locations of the shipworm. The caecum is devoid of microbes (Betcher, 2011) and siderophores may play a role in nutrient deprivation, as found in plant models (Kloepper et al., 1980). By sequestering the iron in the caecum, siderophores can make iron unavailable to competing microbes, thereby their restricting growth. MS-based screening of dissected organs and tissues of shipworms samples would test this theory. However, due to their small size, precise dissection of *L. pedicellatus* samples was difficult without cross-tissue contamination. Larger shipworm species such as *Bankia setacea* would be more amenable to such studies. In addition, chemical localization studies using methods such as MS-imaging would allow the detection of the location of the siderophore within host tissue. These types of studies would be valuable in determining the extent to which siderophores are utilized in the host and lend insight into their possible roles, such as intersymbiont competition (Chapter 4) or suppressing competing microbes in the caecum.

The novel catecholate siderophore, turnerbactin, has been characterized and the biosynthesis has been analyzed by bioinformatic methods. Iron within the host may be extremely limiting, and siderophore production may be relied upon for iron acquisition. Support for its role in the symbiosis is provided by the detection of turnerbactin in shipworm samples. Figure 2.13 presents the proposed role of various proteins encoded in the biosynthetic pathway in the acquisition of iron.

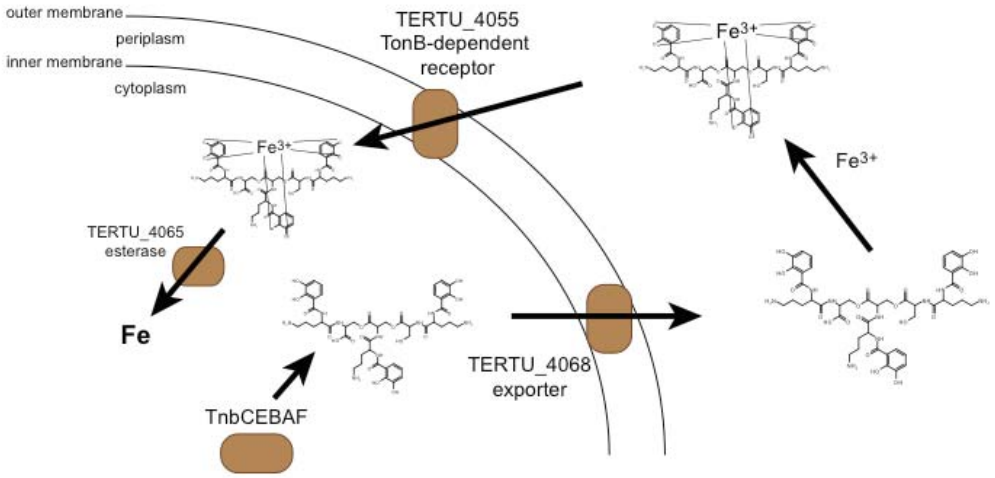

Figure 2.13. Proposed role of proteins in the biosynthetic pathway in the siderophore-mediated acquisition of iron.

3 A profile of siderophores from shipworm symbionts

3.1 Introduction

Although *Teredinibacter turnerae* has been found in numerous genera and species of teredinid bivalves from across the globe (Distel et al., 2002b), it has been estimated to make up only about 10% of the bacterial community in the shipworm *Lyrodus pedicellatus* (Luyten et al., 2006). Shipworms actually harbor a consortium of endosymbionts. This is in contrast to many other bivalve-endosymbiont associations in which a single endosymbiont is detected, as reported by Eisen et al. (1992), Distel et al. (1988), and Distel and Cavanaugh (1994).

Analysis of the 16S rRNA gene sequences from the community of shipworm symbionts reveals a group of closely related, phylogenetically distinct gammaproteobacteria (Distel et al., 2002a; Luyten et al., 2006; Betcher, 2011). Furthermore, within this closely related group, these studies have found that sequences group into discrete clusters.

This work presents a profile of the siderophores from shipworm symbionts. A phylogenetic reconstruction of 16S rRNA gene sequences from shipworm symbionts is presented. This analysis reveals the clustering of symbiont sequences into discrete clades, supporting previous results. One clade was determined to be the *T. turnerae* clade, comprised of multiple strains of *T. turnerae*. A survey of a number of different *T. turnerae* strains was conducted to determine the extent of turnerbactin's prevalence within this clade. Also, representative sequences from other clades of shipworm symbionts were examined for the potential for siderophore biosynthesis through

genomic analysis. These results suggest that catecholate-type siderophores predominate in shipworm symbionts.

3.2 Methods

3.2.1 Phylogenetic analysis of shipworm symbionts

Sequences of shipworm symbionts were provided by Meghan Betcher and collaborators in the ICBG-PMS group. Sequences were aligned using ClustalX (Larkin et al., 2007) and edited in Geneious v5.4 (Drummond et al., 2011). Hypervariable regions were excluded from the analysis. Phylogenetic reconstruction was performed with RAxML v. 7.2.8 using the General Time Reversible model of nucleotide substitution (Tavare, 1986) under the Γ model of rate heterogeneity (GTRGAMMA) with 100 bootstrap replicates. The selected tree topology had the highest likelihood score out of 100 heuristic tree searches, each search beginning with a distinct randomized maximum parsimony starting tree. This analysis was done using the facilities made available by the CIPRES portal in San Diego, CA (Miller et al., 2010). A tree using the neighbor-joining method using the Jukes-Cantor genetic distance model (Jukes and Cantor, 1969) was also constructed using the tree building plugin found in Geneious v5.4.

3.2.2 Detection of siderophores from bacterial cultures

Various strains of *T. turnerae* were grown in a low-iron, modified version of shipworm basal medium (SBM) (Distel et al., 2002b; Waterbury et al., 1983) containing 750 ml artificial seawater (Kester et al., 1967), 250 ml distilled water, 0.1

mM KH_2PO_4, 0.094 mM Na_2CO_3, 0.01 mM $Na_2MoO_4.2H_2O$, 0.5% (w:v) sucrose, 5 mM NH_4Cl, 20 mM HEPES buffer (pH 8.0), 0.1 uM EDTA-chelated ferric iron (Sigma), and 1 ml A5+Co trace metal mix (Rippka et al., 1979). Cultures were grown in 125 mL flasks with 50 mL culture medium at 30° C on an orbital shaker at 120 rpm. Cultures were grown for 1-3 days, monitored twice a day for siderophore production with the chrome azurol sulfonate (CAS) assay (Schwyn and Neiland, 1987). Cultures were harvested when CAS activity was detected in culture supernatants. Strains 1162T.0a.2, Bs02, and Bs08 were also grown in iron-deficient SBM for siderophore characterization.

A crude siderophore extract was collected from each strain of *T. turnerae* by separating cells from supernatant using centrifugation at 10000 rpm for 25 min. Decanted supernatant was placed in a new 125 mL flask, ~3 g Diaion HP20 resin was added to the flasks, and the flasks were then incubated at 4° C on an orbital shaker at 110 rpm for 4 hrs. The mixture was poured into a column allowing the spent supernatant to flow through. The HP20 resin was washed with MilliQ water and then the siderophores were eluted with 25% isopropanol, detected by the CAS assay. The 25% isopropanol fraction was concentrated by rotary evaporation *in vacuo*.

The concentrated siderophore HP20 extract was directly injected onto a ThermoElectron LTQ-Orbitrap high resolution mass spectrometer. Turnerbactins were detected using positive mode high resolution electrospray ionization mass spectrometry (HRESIMS) and tandem MS. Selected ion monitoring (SIM) was used with HRESIMS, set to monitor the mass range 355.6-356.6, a range which brackets

the m/z 356.1 peak of the molecular ion of turnerbactin. Purified turnerbactin was used as a standard.

3.2.3 Bioinformatic analysis

Module identification and domain organization of the NRPS was carried out by the online tools NRPS-PKS (Ansari et al., 2004) and the PKS/NRPS analysis website (Bachmann and Ravel, 2009).

Adenylation domain: The specificity-conferring code of A domains were determined using three different online tools: NRPS-PKS (Ansari et al., 2004), NRPSpredictor2 (Rottig et al., 2011), and the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Phylogenetic analysis of the specificity-conferring codes of A domains was also carried out. Protein sequences of A domains from the NRPS's putatively involved in siderophore production in Bs02, Bs08, and Bs12 were retrieved using the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Reference A domain protein sequences were compiled from a list of A domains with annotated specificity supplied by Rausch et al. (2005). A multiple sequence alignment was constructed using MUSCLE (Edgar, 2004). After extracting approximately a 100 amino acid stretch between core motifs A4 and A5, the alignment was manually edited, guided by the structural anchors of this region identified by Stachelhaus et al. (1999) and Challis et al. (2005). The alignment was manipulated and edited in Geneious v5.4 (Drummond et al., 2011). The eight residue alignment was exported out of Geneious and into MEGA5 (Tamura et al., 2011). Phylogenetic trees were inferred using the unweighted pair group method arithmetic

averages (UPGMA) (Sneath and Sokal, 1973) with the JTT (Jones et al., 1992) amino acid substitution matrix. The rate variation among sites was modeled with a gamma distribution.

Condensation domain: Protein sequences of condensation domains were retrieved using the PKS/NRPS analysis website (Bachmann and Ravel, 2009). Aligned protein sequences of condensation domains were downloaded from Rausch et al. (2007). The condensation domains from the NRPS's identified to be putatively involved in siderophore production in Bs02, Bs08, and Bs12 were added to the downloaded condensation domain alignment using the multiple sequence alignment program MUSCLE. The alignment was edited in Geneious v5.4. A maximum likelihood (ML) tree of condensation domains was reconstructed using RAxML 7.2.8 (Stamatakis, 2006). The amino acid substitution matrix used in this analysis was the JTT matrix (Jones et al., 1992), with the Γ model of rate heterogeneity. RAxML's rapid bootstrap was performed with 100 replicates and the best scoring ML tree was saved. This analysis was done using the facilities made available by the CIPRES portal in San Diego, CA (Miller et al., 2010).

TE domain: The sequence of the TE domain was aligned with other annotated NRPS TE domains using MUSCLE. The alignment was edited in Geneious v5.4. The curation of the alignment was guided by the alignments presented in the work by Bruner et al. (2002) and Samel et al. (2006), who reported the structures of SrfTE and FenTE, respectively.

NRPS-independent siderophore biosynthetic gene clusters were searched for using profile Hidden Markov Models (HMM), statistical models extracted from

multiple sequence alignments, and the software package HMMer (Eddy, 1998) available at http://hmmer.org. The profile HMM of NRPS-independent siderophore biosynthetic genes (PF04183) was downloaded from the Pfam database and HMMer was used to search selected genomes for the presence of these genes.

3.3 Results

3.3.1 Phylogenetic analysis of symbiont strains

A maximum likelihood phylogenetic reconstruction of shipworm symbiont strains reveals the clustering of symbiont strains into discrete clades (Figure 3.1). Similar results were found by Distel et al. (2002a) and Betcher (2011). Clades were numbered in such a way as to maintain a consistent numbering scheme with Betcher (2011). Bootstrap resampling was performed to determine the statistical significance of these clades. The relatively high bootstrap values at the basal branch of each clade support the observed clustering of symbionts according to 16S rRNA phylogeny. A phylogenetic tree was also reconstructed using the neighbor-joining method and supports the clustering pattern observed in the maximum likelihood analysis (data not shown).

As reported by Betcher (2011), the clustering of symbionts can be host-specific. Specifically, all strains isolated from the shipworm *Bankia setacea*, named with the prefix Bs, group together in Clades 3, 4, and 5. No strains from *B. setacea* group with any of the *T. turnerae* strains (Clade 1), consistent with the finding that *T. turnerae* has never been found in *B. setacea* samples (D. Distel, personal communication).

3.3.2 MS survey of *Teredinibacter turnerae* strains

A survey of *T. turnerae* strains was carried out to determine the extent of turnerbactins in Clade 1 of shipworm symbionts. Strains of *T. turnerae* were grown in low-iron conditions to induce siderophore production and a HP20 crude extract of culture supernatants was prepared for each strain. The crude extract was then directly screened for the presence of turnerbactin using HRESIMS in the SIM mode set to scan a one mass unit range centered on m/z 356.1, the $[M+H]^+$ of turnerbactin. If an m/z peak was detected matching that of turnerbactin, tandem MS was also performed to compare the fragmentation pattern of the compound with that of turnerbactin. Tandem MS of turnerbactin yields three major m/z peaks: 338.2, 251.1, and 115.1. A positive result for the presence of turnerbactin includes a matching high-resolution m/z peak and a matching fragmentation pattern. The results of this analysis are shown in Table 3.1. All tested strains of *T. turnerae* showed the presence of turnerbactin. This result suggests that the turnerbactins are ubiquitous in Clade 1 of the shipworm symbionts. The host of the strains from which turnerbactin was detected span multiple genera and species of shipworms.

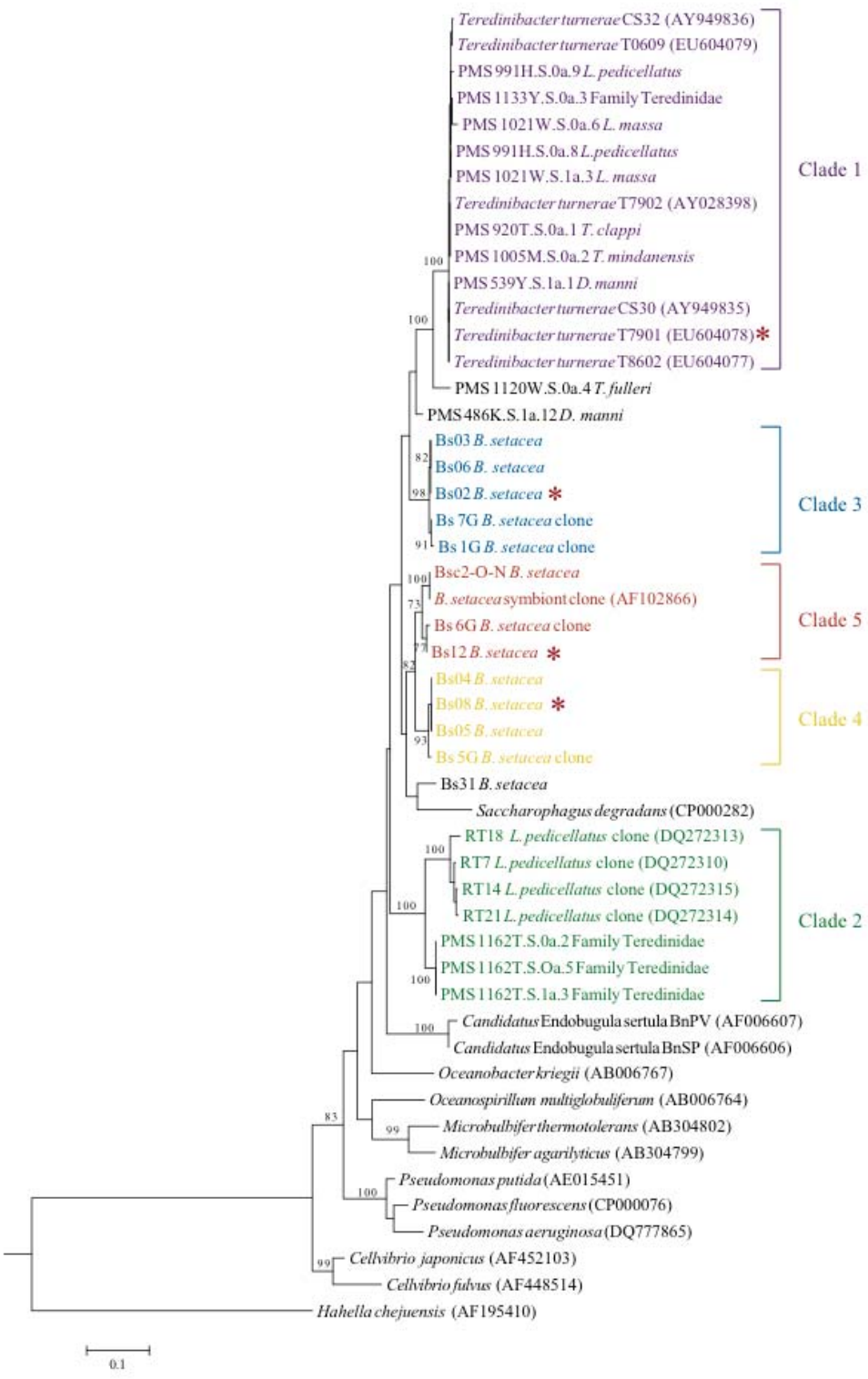

Figure 3.1. Maximum likelihood tree of partial 16S rRNA gene sequences from shipworm symbiont strains, showing the clustering of symbiont strain sequences into discrete clades. A strain without an accession number represents a sequence provided by the PMS-ICBG. * indicates a shipworm symbiont strain whose genome has been sequenced. Boostrap values over 70% are shown and are based on 100 replicates. Scale bar represents 0.1 nucleotide substitutions per site.

Table 3.1. List of *T. turnerae* strains, the respective host from which the strain was isolated, and the result of MS/MS analysis of the presence of turnerbactin in the supernatants of iron-deficient cultures.

Strain	Host	Presence of turnerbactin by MS/MS (+/-)
T7901	*Bankia gouldi*	+
T7902	*Lyrodus pedicellatus*	+
T7903	*Teredo navalis*	+
T8402	*Teredora malleolus*	+
T8602	*Dicyathifer manni*	+
T8201	*Psiloteredo healdi*	+
T8302	*Teredo furcifera*	+
T8303	*Teredo bartschi*	+
CS30	*Neoteredo reynei*	+

3.3.3 Bioinformatic survey of shipworm symbiont genomes

In addition to *T. turnerae* T7901, the genomes of three other shipworm symbiont strains have been sequenced (draft status): Bs02, Bs08, and Bs12. These genomes were screened for the presence of siderophore biosynthetic gene clusters and detected gene clusters were analyzed for the type of siderophore produced.

3.3.3.1 Bs08. Bs08 contains a putative catecholate siderophore biosynthesis pathway (Figure 3.2), evidenced by the presence of genes with homology to the *entCEBA* genes responsible for the production and activation of 2,3-dihydroxybenzoate (DHB) in the biosynthetic pathway of enterobactin (Walsh et al., 1990). Table 3.2 shows proteins with closest similarity to the gene products of this cluster, many of which are from *T. turnerae* T7901. Bs08orf2620 codes for a 2399 amino acid, two module NRPS, similar to that of *T. turnerae* T7901. Bs08orf2620 and *tnbF* share 58% identity and 72% similarity. Just as in *tnbF*, each module in Bs08orf2620 has a C, A, and PCP domain, followed by a C-terminal TE domain.

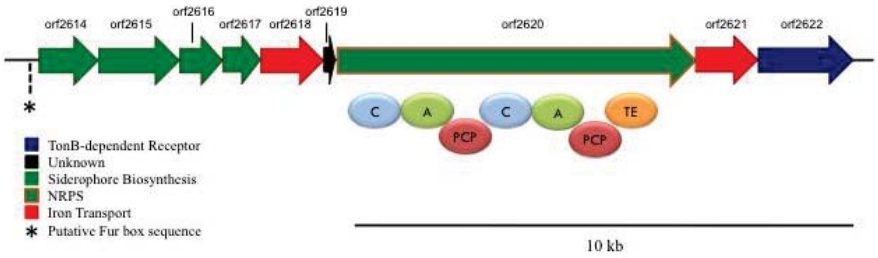

Figure 3.2. Organization of Bs08's putative siderophore biosynthetic gene cluster, with the putative domain organization of the NRPS modules.

Table 3.2. Proteins with similarity to the products of the putative Bs08 siderophore biosynthetic gene cluster. S: similarity, I: identity.

Locus	Protein size	Similarity	S/I (%/%)
orf2614	394	isochorismate synthase DhbC (*Teredinibacter turnerae* T7901)	74/61
orf2615	544	Enterobactin synthetase component E/2,3-dihydroxybenzoate-AMP ligase (*Teredinibacter turnerae* T7901)	77/65
orf2516	289	isochorismatase (*Teredinibacter turnerae* T7901)	78/66
orf2617	255	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (*Teredinibacter turnerae* T7901)	77/64
orf2618	423	Enterochelin esterase (*Yersinia intermedia*)	57/43
orf2619	76	MbtH-like protein (*Teredinibacter turnerae* T7901)	69/51
orf2620	2399	Nonribosomal peptide synthetase TERTU_4067 (Teredinibacter turnerae T7901)	72/58
orf2621	420	putative enterobactin exporter EntS (*Teredinibacter turnerae* T7901)	78/66
orf2622	644	TonB-dependent receptor TERTU_4055 (*Teredinibacter turnerae* T7901)	81/66

This gene cluster also shows a putative Fur box sequence (Figure 3.2), suggesting transcriptional regulation by one of the three Fur homologs found in the genome. This sequence is compared to the bacillibactin biosynthetic operon of *Bacillus subtilis* in Table 3.3, whose Fur box matches exactly to the classic Fur box sequence proposed by de Lorenzo et al. (1987), the Fur box sequences from the turnerbactin biosynthetic gene cluster are also shown.

Table 3.3. Putative Fur box sequences of Bs08 and Bs12's putative siderophore biosynthetic gene cluster compared to the bacillibactin biosynthetic operon of *Bacillus subtilis* and the putative Fur box sequences of turnerbactin's biosynthetic operon. Shaded bp's match the revised view of Fur box sequences consisting of two overlapping 7-1-7 motifs proposed by Baichoo and Helmann (2002).

	7-1-7 consensus sites TGATAAT-ATTATCA TGATAAT-ATTATCA (L) TGATAAT-ATTATCA (R)
Operon	
dhbABCKF	ATTGATAATGATAATCATTATCAATAGATTG
TERTU 4055-58	AATGATAATGCGAATAATTATCATATAGATT
TERTU 4059-68	CACAATAATGCGAATCATTCTCATTATTGA
Bs08orf2614-2622	GATACAGTTGCGAACCGTTATCATTAGCATT
Bs12or2083-2074	ACTCTGATTGAGAATTGTTCTCAGTTGCATT

Bioinformatic analyses were carried out to predict the peptide product from this gene cluster. Three online tools were used to identify the specificity-conferring code of Bs08orf2620's A domains and predict specificity: NRPS-PKS (Ansari et al., 2004), the NRPSpredictor2 (Rottig et al., 2011), and the PKS/NRPS analysis website (Bachmann and Ravel, 2009). The three online tools predicted the same specificity for the A domain of the second module (M2) of Bs08orf2620 (Table 3.4). However, the prediction for the A domain of first module (M1) was inconclusive. The NRPS-PKS tool and PKS/NRPS analysis website reported a similar specificity-conferring code for this A domain, while the NRPSpredictor2 reported a different code. In all cases, the specificity codes predicted for Bs08orf2620 were very similar to those predicted for *T. turnerae*'s *tnbF*. Much like *tnbF*, all three programs failed to identify a reliably significant prediction for the M1 A domain of Bs08orf2620.

Table 3.4. Specificity codes of Bs08orf2620's A domains predicted by three separate online tools. The specificity codes from *tnbF*'s A domain are shown for comparison. (-) indicates an undetermined residue at that particular position.

	Prediction tool

	NRPS-PKS	PKS/NRPS analysis website	NRPSpredictor2
Bs08orf2620 M1 A domain	DSWDIILV	DSWDII--	DSADAGLV
tnbF M1 A domain	D-WDIILV	D-WDIILV	DSDDGGLV
Bs08orf2620 M2 A domain	DVWHFSLV	DVWHFSLV	DVWHFSLV
tnbF M2 A domain	DVWHFSLV	DVWHFSLV	DVWHFSLV

To further analyze the specificity prediction of Bs08orf2620's A domains, a phylogenetic analysis of the specificity-conferring codes of Bs08orf2620's A domains was carried out with a selection of A domains with annotated specificity (Figure 3.3). The results of the analysis show that the M2 A domain groups with other serine-activating A domains, as expected. The M1 A domain groups tightly with the M1 A domain from *tnbF,* with other lysine-activating A domains. The results of the A domain specificity predictions suggest that the same amino acids are activated by Bs08orf2620 as *tnbF*.

Rausch et al. (2007) demonstrated that the reconstructed phylogeny of condensation domains shows a functional grouping of C domains. Phylogenetic analysis of the C domains was carried out to determine the functional type of C domains present in Bs08orf2620. As expected from its location in the NRPS, the M1 C domain groups strongly with the Starter condensation domains. The M2 C domain of Bs08orf2620 grouped within the $^{D}C_{L}$ condensation domains, grouping closely with the M2 C domain from *tnbF* (Figure 3.4). This placement is strongly supported by

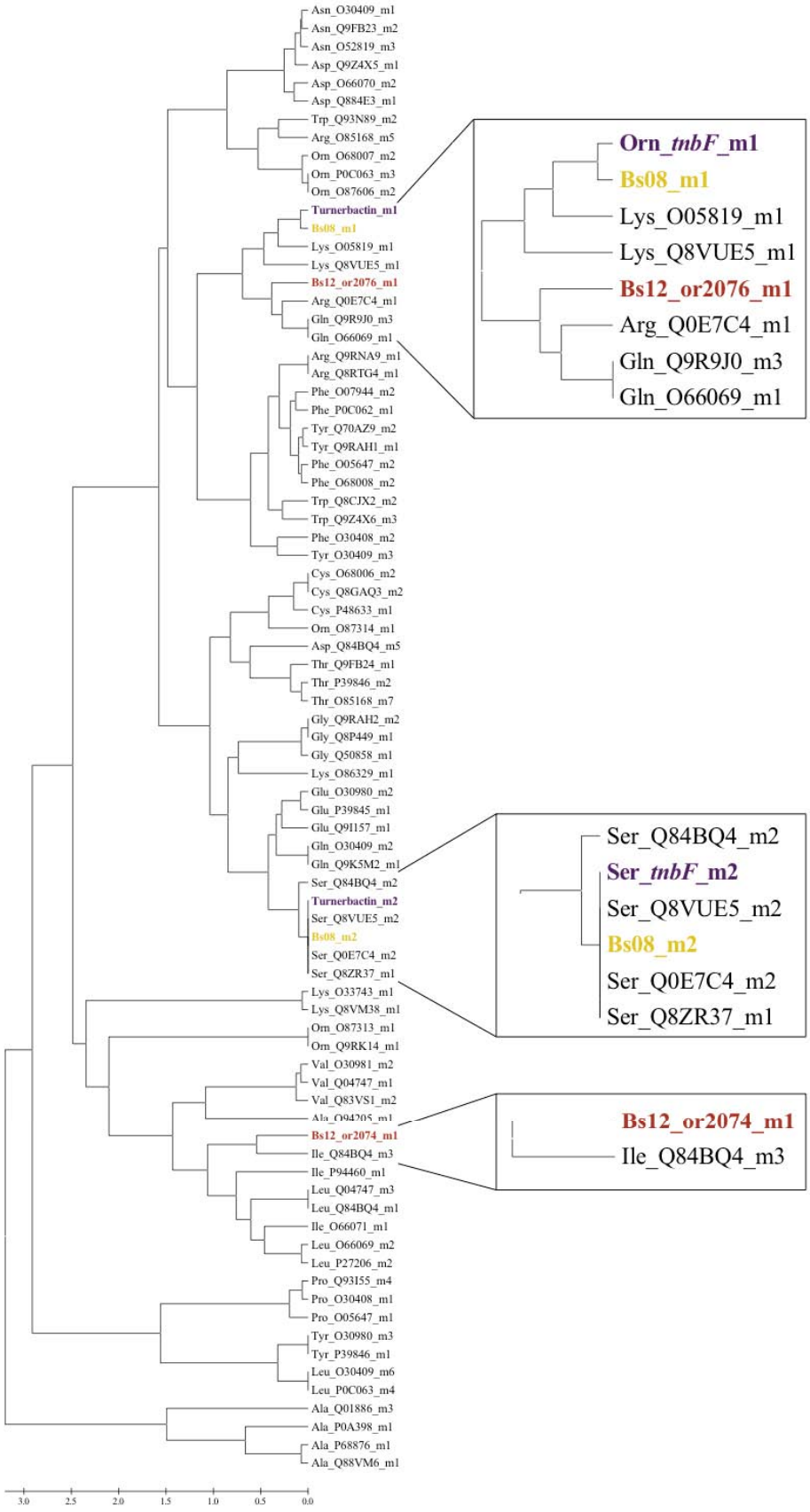
Asn_O30409_m1
Asn_Q9FB23_m2
Asn_O52819_m3
Asp_Q9Z4X5_m1
Asp_O66070_m2
Asp_Q884E3_m1
Trp_Q93N89_m2
Arg_O85168_m5
Orn_O68007_m2
Orn_P0C063_m3
Orn_O87606_m2
Turnerbactin_m1
Bs08_m1
Lys_O05819_m1
Lys_Q8VUE5_m1
Bs12_or2076_m1
Arg_Q0E7C4_m1
Gln_Q9R9J0_m3
Gln_O66069_m1
Arg_Q9RNA9_m1
Arg_Q8RTG4_m1
Phe_O07944_m2
Phe_P0C062_m1
Tyr_Q70AZ9_m2
Tyr_Q9RAH1_m1
Phe_O05647_m2
Phe_O68008_m2
Trp_Q8CJX2_m2
Trp_Q9Z4X6_m3
Phe_O30408_m2
Tyr_O30409_m3
Cys_O68006_m2
Cys_Q8GAQ3_m2
Cys_P48633_m1
Orn_O87314_m1
Asp_Q84BQ4_m5
Thr_Q9FB24_m1
Thr_P39846_m2
Thr_O85168_m7
Gly_Q9RAH2_m2
Gly_Q8P449_m1
Gly_Q50858_m1
Lys_O86329_m1
Glu_O30980_m2
Glu_P39845_m1
Glu_Q9I157_m1
Gln_O30409_m2
Gln_Q9K5M2_m1
Ser_Q84BQ4_m2
Turnerbactin_m2
Ser_Q8VUE5_m2
Bs08_m2
Ser_Q0E7C4_m2
Ser_Q8ZR37_m1
Lys_O33743_m1
Lys_Q8VM38_m1
Orn_O87313_m1
Orn_Q9RK14_m1
Val_O30981_m2
Val_Q04747_m1
Val_Q83VS1_m2
Ala_Q94205_m1
Bs12_or2074_m1
Ile_Q84BQ4_m3
Ile_P94460_m1
Leu_Q04747_m3
Leu_Q84BQ4_m1
Ile_O66071_m1
Leu_O66069_m2
Leu_P27206_m2
Pro_Q93I55_m4
Pro_O30408_m1
Pro_O05647_m1
Tyr_O30980_m3
Tyr_P39846_m1
Leu_O30409_m6
Leu_P0C063_m4
Ala_Q01886_m3
Ala_P0A398_m1
Ala_P68876_m1
Ala_Q88VM6_m1
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Orn_tnbF_m1
Bs08_m1
Lys_O05819_m1
Lys_Q8VUE5_m1
Bs12_or2076_m1
Arg_Q0E7C4_m1
Gln_Q9R9J0_m3
Gln_O66069_m1
Ser_Q84BQ4_m2
Ser_tnbF_m2
Ser_Q8VUE5_m2
Bs08_m2
Ser_Q0E7C4_m2
Ser_Q8ZR37_m1
Bs12_or2074_m1
Ile_Q84BQ4_m3

Figure 3.3. Phylogenetic clustering of A domains using the unweighted pair group method using arithmetic averages (UPGMA). The A domains from Bs08 are shown in yellow, Bs12 are shown in orange, *tnbF* are shown in purple. Each A domain is labeled with its amino acid specificity followed by its UniProtKB accession number followed by the module number of the NRPS from which the A domain is referring.

the high bootstrap values at the base of the Dual E/C and $^{D}C_{L}$ groups. Similar to *tnbF*, no epimerase domain was detected in Bs08orf2620 through pHMM analysis, nor was an external racemase coded for in the biosynthetic gene cluster. This suggests that the M2 C domain of Bs08orf2620, like the M2 C domain of *tnbF*, has $^{L}C_{L}$ activity as opposed to its predicted $^{D}C_{L}$ activity.

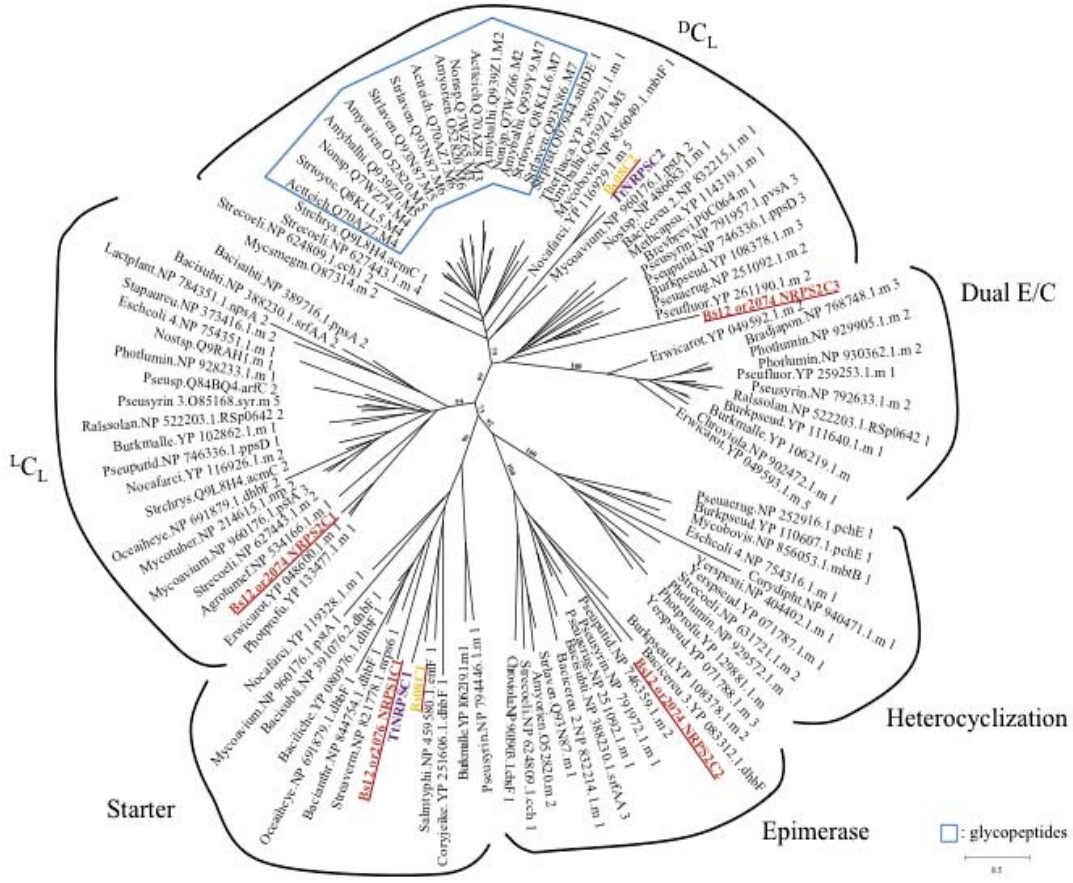

Figure 3.4. Maximum likelihood tree of C domains, showing the grouping of different C domain subtypes. The C domains of from the current study are underlined in red. Each C domain is labeled with the organism name, followed by accession number, followed by the module number from which the C domain is referring. Bootstrap values are based on 100 replicates and are only shown for the basal branches of groups.

A multiple sequence alignment of the TE domain from Bs08orf2620 with other annotated TE domains was constructed to identify the residue in the conserved Pro26 position, shown to influence peptide release from the TE domain of the surfactin biosynthesis NRPS (Tseng et al., 2002). The alignment shows a Pro residue in this conserved position, similar to *tnbF* and other cyclizing TE domains (Figure 3.5).

The results from the bioinformatic analyses of the catalytic domains of Bs08orf2620 suggest that the siderophore produced by Bs08 is very similar, if not exactly the same, as *T. turnerae* T7901.

Figure 3.5. Partial alignment of TE domains. TE domain sequences were extracted from annotated NRPS's. UniProtKB accession numbers of NRPS's are as follows: fengycin, L42523; surfactin, Q08787; gramicidin, P0C064; bacillibactin, P45745; enterobactin, P11454; chrysobactin, Q8VUE5; vanchrobactin, Q0E7C4; pyochelin, O85740; yersiniabactin, Q9Z373. The residue highlighted in red represents the conserved Pro26 in SrfTE, hypothesized to indicate an intramolecular cyclization mechanism of release by TE domains.

3.3.3.2 Bs12. Bs12 also contains a putative catecholate siderophore biosynthetic gene cluster (Figure 3.6). Table 3.5 shows proteins with closest homology to the gene products of this cluster. Homologues of *entCEBA* genes are paired with two genes encoding for NRPS's. The first NRPS, Bs12orf2076, is a 1055 amino acid, one module NRPS, while Bs12orf2074 is a 2080 amino acid, two module

NRPS. It's expected that the *entCEBA*-like genes and the two NRPS's work in an assembly line fashion to produce one peptide product.

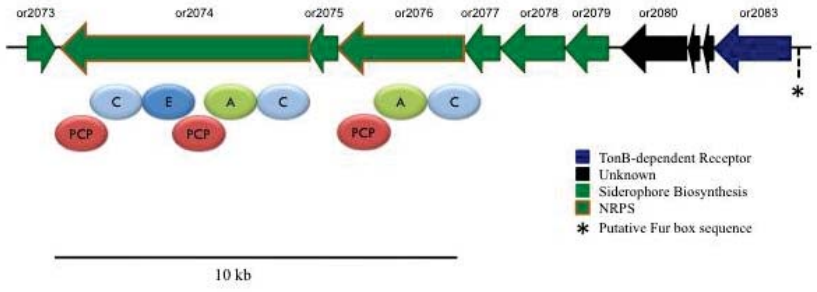

Figure 3.6. Organization of Bs12's putative siderophore biosynthetic gene cluster, with the putative domain organization of the NRPS modules.

Table 3.5. Proteins with similarity to the products of the putative Bs12 siderophore biosynthetic gene cluster. S: similarity, I: identity.

Locus	Protein size	Similarity	S/I (%/%)
or2073	239	Phosphopantetheinyl transferase component of siderophore synthetase (*Hahella chejuensis*)	60/42
or2074	2080	putative non-ribosomal peptide synthetase (*Pectobacterium carotovorum*)	63/46
or2075	250	2,3-dihydroxybenzoate-2,3-dehydrogenase (*Pectobacterium carotovorum*)	67/53
or2076	1055	enterobactin synthetase component F (*Pectobacterium atrosepticum*)	66/49
or2077	305	Isochorismatase (*Pectobacterium carotovorum*)	72/56
or2078	543	enterobactin synthase subunit E (*Halomonas elongata*)	72/57
or2079	411	isochorismate synthase (*Photobacterium profundum*)	66/45
or2080	554	PepSY-associated TM helix domain protein (*Teredinibacter turnerae* T7901)	61/42
or2081	101	hypothetical protein (*Shewanella* sp. HN-41)	67/51
or2082	89	hypothetical protein (*Shewanella* sp. HN-41)	56/35
or2083	650	Outer membrane protein (*Oceanospirillum* sp. MED92)	53/35

This gene cluster also shows a putative Fur box sequence (Figure 3.6), suggesting transcriptional regulation by one of the three Fur homologs found in the genome. This sequence is compared to the bacillibactin biosynthetic operon of *Bacillus subtilis* in Table 3.3, whose Fur box matches exactly to the classic Fur box

sequence proposed by de Lorenzo et al. (1987), the Fur box sequences from the turnerbactin biosynthetic gene cluster are also shown.

There are two atypical aspects to the NRPS, Bs12orf2074. The first atypical feature is the presence of only one A domain for its two modules. This type of organization has been reported in other NRPS systems, such as for the siderophore yersiniabactin, whose biosynthetic machinery includes an NRPS with three modules but only one A domain. Yersiniabactin's sole A domain activates and supplies a Cys residue to the PCP domain of each module (Keating et al., 2000). A similar mode of biosynthesis is expected to be carried out by Bs12orf2074. The second feature of Bs12orf2074 that is unique is that there is no detected TE domain. Aside from TE domain peptide release, other chain release mechanisms known for NRPS's include a C-terminal reductase domain (Re domain), leading to an aldehyde product, or a C-terminal C domain that release C-terminal amides (Keating et al., 2001). However, these alternative peptide release domains were not detected in Bs12orf2074. It is possible that an external thioesterase acts to release the peptide product from the NRPS, but a lone thioesterase domain was not found in this particular gene cluster.

Bioinformatic analyses were carried out to predict the peptide product from this gene cluster. All three online A domain analysis tools report a specificity-conferring code of DAEDIGTV for Bs12orf2076. NRPS-PKS predicts a specificity for Lys, the NRPSpredictor2 predicts a specificity for Glu, and the PKS/NRPS analysis website predicts a specificity for Arg. While the predicted amino acid varies between the three online tools, they commonly predict for a hydrophilic amino acid. The specificity-conferring code reported by the three online A domain analysis tools

for Bs12orf2074 is DAMVGGCV, however, these tools were not able to predict a consensus specificity of the A domain.

Phylogenetic analysis of the A domains from NRPS's of Bs12's putative siderophore biosynthetic gene cluster was carried out to further analyze the predicted specificity of its A domains. This analysis shows the clustering of Bs12orf2076's A domain with an Arg-specific A domain, while the A domain of Bs12orf2074 groups with Ile-specific A domains (Figure 3.3).

Phylogenetic analysis of the C domains was carried out to determine the functional type of C domains present in Bs12's two NRPS's (Figure 3.4). Also included in the analysis was the detected E domain of Bs12orf2074. Unlike the results from *tnbF* and Bs08's siderophore NRPS C domain analysis, the C domains from Bs12's siderophore biosynthetic gene cluster group as expected according to the observed NRPS domain organization. Bs12orf2076's sole C domain groups strongly with the Starter condensation domains, similar to the M1 C domains from *tnbF* and Bs08orf2620. The M1 C domain from Bs12orf2074 groups with $^{L}C_{L}$ functional domains as expected due to the absence of an E domain in the upstream module. The M2 C domain groups with other E domains as expected from its annotation as an E domain. And the M3 C domain groups with $^{D}C_{L}$ functional domain as expected from the presence of an E domain immediately upstream.

3.3.3.3 Bs02. While Bs02 does code for putative secondary metabolite gene clusters (data not shown), no siderophore biosynthetic gene cluster was identified through this work. However, it should be noted that the genome is in draft status and may not accurately reflect its true siderophore biosynthetic capability.

3.3.4 Growth experiments to induce siderophore production

Available representative strains (outside of Clade 1) were grown in iron-deficient SBM to induce siderophore production for isolation and characterization. The following strains were grown: 1162T.0a.2 (Clade 2), Bs02 (Clade 3), and Bs08 (Clade 4). All three strains failed to show positive CAS assay results, even for Bs08 whose genome predicts the production of a catecholate siderophore. The particularly poor growth of Bs08 in culture suggests that the current culture conditions are not ideal and need to be optimized for efficient laboratory growth. The strain Bs12 was not available for culture.

3.4 Discussion

This work presents a siderophore profile of shipworm symbionts through MS screening and a bioinformatic survey. The sequence and phylogenetic analysis of the 16S rRNA genes of symbiont strains shows a group of closely related bacteria that group into discrete clades. It has been suggested that these clades represent ecological differentiation amongst the community (Luyten et al., 2006). It's believed that these clades are an accumulation of neutral diversity and the strains within each clade occupy the same ecological niche. They persist because competitive mechanisms are too weak to purge diversity from within the clusters (Acinas et al., 2004). Therefore, representatives from each clade were chosen to represent the entire clade's niche within the shipworm system.

Intraclade neutral diversity was supported by the MS survey of *Teredinibacter turnerae* strains (Clade 1), which revealed the ubiquity of turnerbactin in this species of bacteria. The strains of *T. turnerae* screened in this MS survey originate from a wide range of shipworm hosts. With the finding of turnerbactin in samples of *L. pedicellatus*, these results suggest turnerbactin may play a role in the symbiosis of a wide range of teredinid bivalve mollusks (Figure 3.7).

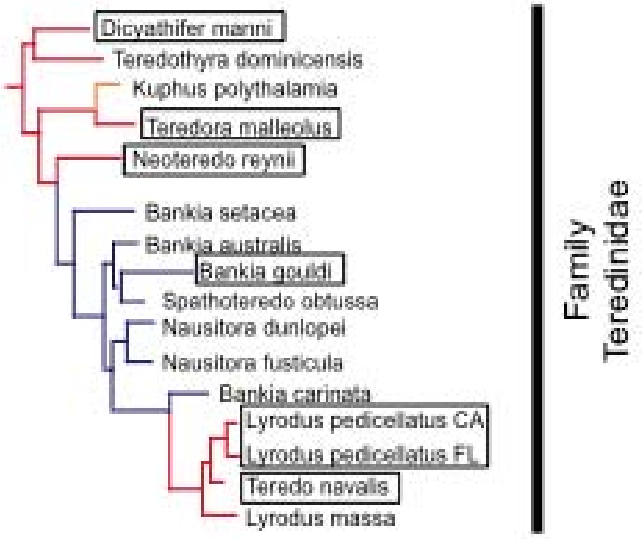

Figure 3.7. Proposed distribution of turnerbactin in shipworms. Phylogram of concatenated small and large subunit nuclear rRNA genes from representative shipworms in the Family Teredinidae. Turnerbactin was detected in strains of *T. turnerae* isolated from boxed shipworm species. Figure adapted from Distel et al. (2011).

The bioinformatic survey of three genome-sequenced strains from three different shipworm symbiont clades provides an expanded view of the potential for siderophores from shipworm symbionts. Representatives from Clades 4 and 5 show the potential to produce catecholate siderophores. Analysis of Bs08 (Clade 4) reveals

the potential for a siderophore very similar to turnerbactin, while Bs12 (Clade 5) reveals the potential for a catecholate siderophore most likely different in structure to that of turnerbactin. The results of the siderophore survey are shown in Table 3.6, along with relevant functional attributes of each clade. The results show the potential for production of catecholate-type siderophores exclusively from shipworm symbionts examined to this point. Due to the pK_a of binding groups, catecholate siderophores are able to outcompete other functional type of siderophores at physiological pH, such as carboxylate-type siderophores (Miethke and Marahiel, 2007). A variety of other host-associated microbes (mostly pathogens) utilize catecholate-type siderophores in their respective associations. The archetypical catecholate siderophore, enterobactin, is utilized by a variety of enteric bacteria from *Escherichia coli* (O'Brien and Gibson, 1970) to *Salmonella typhimurium* (Pollack and Neilands, 1970). The fish pathogen *Vibrio anguillarum* relies on a plasmid-encoded iron-acquisition system and the ability to produce the catecholate siderophore anguibactin for virulence (Crosa, 1980; Crosa et al., 1980; Wolf and Crosa, 1986). Similarly, the ability to produce the catecholate siderophore vulnibactin by *Vibrio vulnificus* affects its virulence (Litwin et al., 1996).

Table 3.6. Summary of the symbiosis-contributing phenotypes and the siderophore potential of the five clades of shipworm symbionts. A (+) cellulase production indicates the demonstrated ability to degrade cellulose in laboratory cultures. ND: not detected.

	Clade 1 (*Teredinibacter turnerae*)	**Clade 2 (1162T.0a.2)**	**Clade 3 (Bs02)**	**Clade 4 (Bs08)**	**Clade5 (Bs12)**
N_2 fixation	+	-	-	+	-
Cellulase Production	+	+	+	+	+
Siderophore	Turnerbactin	ND	ND	Turnerbactin-like	Catechol siderophore

Phylogenetic analysis of shipworm symbionts shows that *B. setacea* symbiont sequences group separately from symbiont sequences from all other shipworm species compiled in this and previous studies (Betcher, 2011). From the available data, the symbionts from *B. setacea* can thus be treated as a separate community from the rest of the symbionts. Bs08 in *B. setacea* seems to serve an analogous role as *T. turnerae* does in other shipworm hosts, with its ability to fix nitrogen as well as potentially producing a high-affinity siderophore. Growth under diazotrophic conditions has been shown to require more iron than under non-diazotrophic growth conditions (Kustka et al., 2003), therefore, nitrogen fixers such as *T. turnerae* and Bs08 may have a higher demand of iron in the host compared to symbionts such as 1162T.0a.2 and Bs02.

One of the questions driving this siderophore survey of shipworm symbionts is whether the symbionts work together as one cooperative unit in the acquisition of iron or whether they work independently to find a source of iron. Regardless of their ability to fix nitrogen, almost all microorganisms require iron to carry out basic cellular functions. In the host, iron is tightly regulated by host iron-binding proteins (Ratledge and Dover, 2000; Chu et al., 2010). Siderophores represent one route of iron acquisition. The finding that turnerbactin is present in *L. pedicellatus* extracts suggests that the siderophore is utilized in some form in the symbiosis. A structurally diverse set of siderophores from different symbiont clades would suggest competition amongst symbionts for iron, whereas a homogenous set of siderophores would suggest cooperation. The results of the bioinformatic survey of genome-sequenced strains suggest that shipworm symbionts have the potential to produce structurally

different siderophores, as evidenced by Bs12's biosynthetic gene cluster. However, an important step in siderophore-based iron acquisition not addressed in this current survey is the uptake of ferric-siderophore complexes. Siderophores are taken up by specific outer membrane receptors. As a general rule, microbes are able to recognize and take up natively produced siderophores. But examples of microbes to take up heterologous siderophores are numerous (Baig et al., 1986; Simon et al., 1995; Loper and Henkels, 1999; Cornelis and Matthijs, 2002). The ability to take up heterologous siderophores would confer an evolutionary advantage, and this ability has not been studied in shipworm symbionts. Bioassays with shipworm symbionts and their siderophores will reveal the potential for these mechanisms.

4 Role of siderophores in the interaction between symbionts

4.1 Introduction

Shipworms harbor a consortium of closely related endosymbionts. This is in contrast to many other bivalve-endosymbiont associations in which a single endosymbiont is detected, as reported by Eisen et al. (1992), Distel et al. (1988), and Distel and Cavanaugh (1994). Exceptions to the maintenance of a single endosymbiont have been reported. A hydrothermal vent mussel (family Mytilidae) was shown to contain both methanotrophic and thioautotrophic endosymbionts in the same bacteriocyte within the host gill (Distel et al., 1995). Analogously, the gutless marine oligochaete *Olavius algarvensis* contains both sulfate-reducing and sulfide-oxidizing endosymbionts within the same host cell (Dubilier et al., 2001). In the case of the hydrothermal vent mussel and *O. algarvensis*, different symbiont types are found to colonize the same host cell. However, studies investigating the habitation patterns of shipworm symbionts suggest that different groups of symbionts do not inhabit the same host cell. Phylogenetic analyses of shipworm symbionts reveal the clustering of symbionts into discrete clades (Chapter 3; Distel et al., 2002a; Luyten et al., 2006; Betcher, 2011). Localization studies suggest that these clades correlate with a pattern of segregation within the gills (Distel et al., 2002; Betcher, 2011).

The mechanisms of segregation within the host are unclear. However, studies on the microbial dynamics in plants suggest that siderophores might play a role in controlling the growth of associated microbes. Kloepper et al. (1980) showed that plant growth-promoting strains of *Pseudomonas* were able to repress the growth of a plant pathogen. They propose that these plant growth-promoting strains colonize

roots and sequester the limited iron available with siderophores, thereby making the iron unavailable to other microorganisms. This mechanism of action occurs because microbes are generally unable to take up foreign siderophores. Their siderophore uptake receptors recognize and are specific for the siderophores that they produce. Exceptions to this rule are numerous, however (Baig et al., 1986; Simon et al., 1995; Loper and Henkels, 1999; Cornelis and Matthijs, 2002). Species from various genera have shown the ability to take up heterologous siderophores, thereby conferring an evolutionary advantage over those species that are able to only take up native siderophores. Guan et al. (2001) used sponge-associated and seawater bacteria to show the ability of exogenous siderophores to act as growth-promoting and signaling factors. D'Onofrio et al. (2010) were able to culture previously uncultured organisms with the help of exogenous siderophores. Regardless of a beneficial or antagonistic nature, these studies show the potential for siderophores to affect the growth of other members of its community.

The current work aims to understand how siderophores might play a role in either growth-promotion or growth-inhibition of shipworm symbionts. In the previous chapter, phylogenetic clustering of symbionts into clades was presented and a siderophore profile was created to show the prevalence of catecholate siderophores among these clades. However, the cross-utilization potential of these siderophores among symbionts is unknown. By using cross-streaking and disc diffusion bioassays, this work shows that symbiont-derived siderophores are capable of inhibiting the growth of other symbionts, suggesting that there is not universal cross-utilization of siderophores by shipworm symbionts.

4.2 Methods

4.2.1 Bacterial strains and media

Shipworm strains were grown and assayed on a modified version of shipworm basal medium (SBM) (Distel et al., 2002b; Waterbury et al., 1983) containing 750 ml artificial seawater (Kester et al., 1967), 250 ml distilled water, 0.1 mM KH_2PO_4, 0.094 mM Na_2CO_3, 0.01 mM $Na_2MoO_4.2H_2O$, 0.5% (w:v) Sigmacell cellulose Type 101 (Sigma), 5 mM NH_4Cl, 20 mM HEPES buffer (pH 8.0), 10 μM EDTA-chelated ferric iron (Sigma), 1 ml A5+Co trace metal mix (Rippka et al., 1979), and 1.5% agar (w:v) for plates.

4.2.2 Cross-streaking of symbiont strains

The siderophore-production mutant strain of *T. turnerae*, TtAH03, was used in cross-streaking assays to determine the inhibition of this strain on other symbiont strains. The construction of this mutant was described in Chapter 2. A test strain was streaked down the center of a SBM plate. Then TtAH03 was streaked perpendicular to the test strain, as well as wild-type *T. turnerae* as a control (Figure 4.1). The inhibition of growth by the test strain was measured and recorded. The mutant strain contains a chloramphenicol resistance cassette inserted into the M1 C domain of *tnbF* (Figure 2.1, Chapter 2). In order to carry out the cross-streaking experiment, the medium could not contain Cm due to the sensitivity of the test strain. Therefore, the presence of the Cm resistance cassette in TtAH03 was checked by PCR after each experiment.

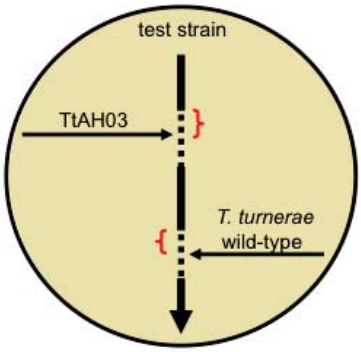

Figure 4.1. Depiction of cross-streaking assay. Red brackets indicate the measured growth inhibition zone of the test strain. Adapted from Brian Fishman.

4.2.3 Disc diffusion bioassays

Symbiont strains representing different shipworm symbiont clades were used in disc diffusion bioassays. Strains used in bioassays were grown in liquid SBM medium overnight and then spread onto an iron-replete SBM plate. Approximately 50 μg of triturnerbactin was applied to a 6-mm-diameter filter disc, which was then placed onto the plate. A filter disc to which water was applied was used as a negative control. The zone of inhibition around the filter disc was measured and recorded.

4.3 Results

4.3.1 Cross-streaking assay

Previous work with shipworm symbionts showed the ability of strains from Clade 1 (*T. turnerae* clade) to inhibit the growth of strains from symbiont clades 2, 3, and 4 (Betcher, 2011). A representative from Clade 5 was not available for the previous study nor for the current study. Additionally, inhibition between any combination of strains from only Clades 2, 3, or 4 was not observed. Recently, a siderophore-producing mutant of *T. turnerae* T7901 was constructed called TtAH03.

The NRPS responsible for siderophore production is disrupted in TtAH03, but this strain maintains wild-type function of the genes responsible for the production of 2,3-dihydroxybenzoate (DHB). This strain was used in cross-streaking assays to determine the role of siderophores in the observed inhibition. The results of these assays are shown in Table 4.1. TtAH03 failed to completely remove inhibition effects toward any of the tested strains. However when compared to cross-streaks with the wild-type, 1162T.0a.2 and Bs08 exhibited a slight reduction in inhibition, while Bs02 did not show any significant change. These results suggest that *T. turnerae* inhibits other symbiont strains by a variety of mechanisms and are not solely the result of siderophore production.

Table 4.1. Inhibition of symbiont strains as a result of cross-streaking and disc diffusion assays. Approximated measured zones of inhibition are in mm.

	Clade 2 (1162T.0a.2)	**Clade 3 (Bs02)**	**Clade 4 (Bs08)**
TtAH03 (cross-streak)	4.5	5	4.75
wild-type (cross-streak)	5	5	5
Disc diffusion with triturnerbactin	1.5	no inhibition	3

4.3.2 Disc diffusion assay

To ascertain the ability of siderophores alone to inhibit the growth of other shipworm symbionts, a disc diffusion assay was carried out with pure triturnerbactin. Triturnerbactin was used because it is most likely to be the native siderophore produced by *tnbF*. Additionally, triturnerbactin is presumed to exhibit the strongest affinity for iron of all the forms of turnerbactin isolated from *T. turnerae* T7901 based

on the calculations by Scarrow et al. (1991) who calculated that the linear trimer of enterobactin has a higher stability constant than that of the DHB-serine dimer. The results of the disc diffusion assays are shown in Table 4.1.

Strain 1162T.Oa.2 (Clade 2) shows inhibition by triturnerbactin. Presumably, triturnerbactin diffuses into the medium and binds the available iron, depriving 1162T.0a.2 of iron immediately surrounding the disc. This suggests that 1162T.0a.2 does not have an outer-membrane receptor that is able to take up the ferric-siderophore complex. This also suggests that 1162T.0a.2 does not produce a siderophore that is able to outcompete triturnerbactin for iron. Siderophore production could not be detected when 1162T.0a.2 was grown alone in iron-deficient media.

Bs02 (Clade 3), on the other hand, does not show any inhibition by triturnerbactin, suggesting that Bs02 has an outer-membrane receptor that is able to take up the ferric-siderophore complex. Alternatively, Bs02 might produce a high-affinity siderophore that is able to outcompete triturnerbactin for iron. However, analysis of Bs02's genome did not reveal any siderophore biosynthetic gene clusters and siderophore production could not be detected when Bs02 was grown alone in iron-deficient media. It should be noted though that this does not preclude the possibility that a siderophore may be produced by Bs02, due to the possibility of an incomplete genome and growth conditions not optimized for this organism.

The most perplexing result from the disc diffusion assays, however, is the inhibition towards Bs08 (Clade 4) by triturnerbactin. Bs08 is predicted to produce a siderophore very similar to that of turnerbactin based on analysis of its putative

siderophore biosynthetic gene cluster. It's possible that Bs08 produces a very similar, yet structurally different siderophore than turnerbactin and its cognate siderophore receptor fails to take up the ferric turnerbactin complex.

4.4 Discussion

This work aimed to determine how siderophores mediate the interactions between shipworm symbionts. Using both cross-streaking and disc diffusion bioassays, heterologous siderophores were shown to affect growth of some symbionts. Cross-streaking experiments with a *T. turnerae* T7901 siderophore-producing mutant shows that this strain is able to inhibit the growth of other symbiont strains, but in the case of 1162T.0a.2 (Clade 2) and Bs08 (Clade 4), the inhibition was slightly decreased compared to wild-type. These results suggest that *T. turnerae* inhibits the growth of other symbionts by multiple mechanisms, presumably by the production of other antimicrobial compounds. This theory is plausible given the presence of multiple putative bioactive metabolite biosynthesis pathways in the genome of *T. turnerae* T7901 (Yang et al., 2009). Disc diffusion assays with triturnerbactin show that the siderophore alone is able to inhibit the growth of symbionts from Clades 2 and 4, presumably through an iron-deprivation mechanism. These results are consistent with the cross-streaking experiments, which showed slightly decreased inhibition by the siderophore-producing mutant.

Interestingly, the growth of Bs08 (Clade 4) was inhibited by triturnerbactin. Bs08's siderophore is predicted to be very similar, if not the same, as turnerbactin. It may be that Bs08's siderophore undergoes post-assembly line tailoring by enzymes

not encoded in the siderophore biosynthetic gene cluster. Alternatively, the two siderophores could be enantiomers, which could result in failed uptake by the heterologous siderophore receptor. This is the case for pyochelin and enantio-pyochelin from *Pseudomonas aeruginosa* and *P. fluorescens*, respectively. The two siderophores are unable to promote growth in the heterologous species, indicating a highly stereospecific siderophore-uptake system (Hoegy et al., 2009). An analogous scenario exists for enterobactin in *E. coli* K-12, where the synthetic enantioenterobactin does not promote growth of *E. coli* mutants (Neilands et al., 1981). However, bioinformatic analysis suggests that Bs08's siderophore has the same stereochemistry as turnerbactin. Isolation and characterization of the siderophore from Bs08 would help to resolve this issue, but as stated previously, attempts to induce siderophore production in iron-deficient media were unsuccessful.

With the grouping of symbiont strains into discrete clades, we can separate the results of this study by shipworm host: those found in *Bankia setacea*, and those found in other shipworm hosts. For the sake of nomenclature moving forward, the "other" group will be referred to as *Lyrodus*, since both Clade 1 and Clade 2 symbionts have been found in this genus of shipworm. First, the case of the *Lyrodus* shipworm host: clearly, we see a demonstrated dominance of *T. turnerae* (Clade 1) over strains such as 1162T.0a.2 (Clade 2), evidenced by both the cross-streaking and disc diffusion assays. This suggests that if there were to be a battle for iron, *T. turnerae* would win out. This would seem to be an advantage for the host: *T. turnerae* is able to provide a source of fixed nitrogen through nitrogen fixation as well as produce cellulolytic enzymes, while 1162T.0a.2 seemingly only provides the latter

(evidenced by cellulose clearing on agar plates and the inability to grow in media lacking a source of fixed nitrogen). Antagonistic interactions between symbionts have also been seen between primary and secondary symbionts in the pea aphid *Acyrthosiphon pisum* (Sakurai et al., 2005). Segregation within the host may be a way to ensure sufficient access to resources, and *in vitro* experiments demonstrate *T. turnerae*'s capacity to create this segregation. Whether the host aids in segregating *T. turnerae* from 1162T.0a.2 or if *T. turnerae* accomplishes this on its own is unknown.

In the case of *B. setacea*, assuming for the moment that Bs08 (Clade 4) produces a turnerbactin-like siderophore and that Bs02 (Clade 3) is able to utilize that siderophore as in the case of triturnerbactin, there does not seem to be this dominance of one clade over another. In fact, it seems that Bs02 is a social cheater. This strain is able to utilize heterologous siderophores without expending the resources to produce one on its own. Moreover, it does not contribute to the supply of fixed nitrogen to the host. Meanwhile, Bs08 fixes nitrogen and produces siderophores. Presumably, both clades provide cellulolytic enzymes to aid in wood digestion. In this case, Bs08 would understandably attempt to segregate itself from strains such as Bs02 who merely leaches off of others, at least in the case of iron. As opposed to the case of *T. turnerae*, cross-streaking experiments suggest that Bs08 is unable to segregate from Bs02 by its own means (Betcher, 2011). The *B. setacea* strain Bs12 was not able to be addressed in this study due to the strain's unavailability. However, this strain contains a putative catecholate siderophore biosynthetic gene cluster whose product may play a role in the dynamics between *B. setacea* symbionts. Fluorescence in situ hybridization studies with *B. setacea* show that a probe specific for Bs02

produces segregated signal from a probe that targets Bs08 and Bs12 (Betcher, 2011). There is not enough evidence to speculate that Bs12's siderophore is involved in this pattern of segregation, but the results from assays with this siderophore would prove interesting.

Figure 4.2 presents a generalized scheme of the siderophore interplay between shipworm symbionts, summarizing the results of the current set of experiments. The figure shows a hypothetical environment in which different types of siderophore-producing and siderophore-utilizing organisms must compete for a limited store of iron. A high-affinity siderophore producing strain (Clade 1) is able to produce and effectively scavenge for iron. This strain and its siderophore are able to restrict the access of iron to other members in the community. Some members (Clade 2) have no means of scavenging for iron and are unable to take up the high-affinity ferric siderophore complex. Other members produce a low-affinity siderophore (possibly Clade 5) that is outcompeted by the high-affinity siderophore. Then there are members of the community who are considered social cheaters (Clade 3). Without producing a siderophore of its own, this member is able to utilize the high-affinity siderophore, thus reaping the benefits of others without exerting comparable effort. In this scenario, organisms representing Clades 1 and 3 would successfully grow in this localized environment, while the growth of organisms representing Clades 2 and 5 would be restricted.

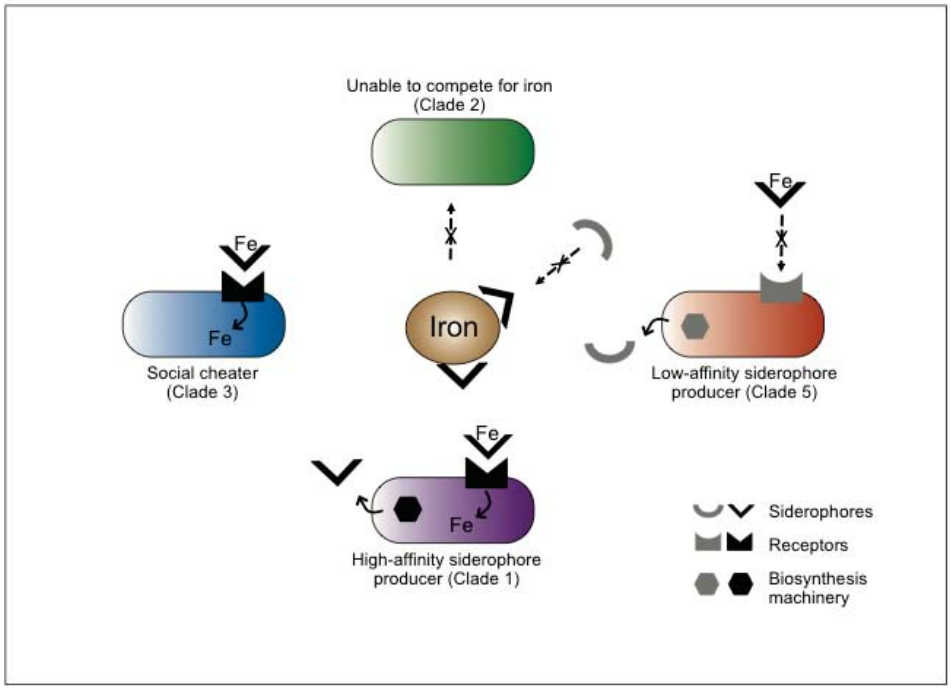

Figure 4.2. Generalized schematic of siderophore interplay between different siderophore-producing and siderophore-utilizing organisms, representing the clades of shipworm symbionts used in the current study. Dotted arrows with X's indicate routes that are blocked. This figure is adapted from Hibbing et al. (2010).

Most likely, the reason why shipworms carry multiple groups of symbionts is that each group provides separate attributes that collectively support the host. Evidence has been shown that the host may penalize symbionts that fail to provide benefit to the host. Work with rhizobia and their legume host showed that symbionts that failed to fix nitrogen were sanctioned by the host by cutting off oxygen supply to non-nitrogen fixing nodules (Kiers et al., 2003). The host imposes a selective environment which favors the stabilization of a mutualistic symbiosis.

Given their attributes discussed above, the different shipworm symbiont groups may each provide a set of cellulolytic enzymes that, when taken all together, allow sufficient digestion of wood by the host. A complementary nutritional role of

symbionts is seen in the cedar aphid *Cinara cedri*. Two different endosymbionts, *Buchnera aphidicola* BCc and *Candidatus* Serratia symbiotica SCc each carry a partial tryptophan biosynthesis pathway when, taken together, produce the essential amino acid tryptophan for the host (Gosalbes et al., 2008). The two endosymbionts, in a necessary metabolic complementation, work together to form a stable symbiotic relationship. Notably, these two endosymbionts occupy separate bacteriocytes within the host (Gómez-Valero et al., 2004). In this example, two different symbionts work in tandem to provide benefit to the entire system, yet still reside in separate compartments.

As alluded to earlier, one advantage of keeping different groups of symbionts segregated within the gill is that cohabitation of a bacteriocyte by different functional symbionts may enhance competition. Competition within the bacteriocyte for space and nutrients would divert attention away from host-beneficial actions such as cellulase production and nitrogen fixation.

Understanding the mode of symbiont transfer from one generation to the next in shipworms would provide insight into how this segregation might occur. One report suggests a vertical mode of transmission in shipworms. Sipe et al. (2000) detected symbiont 16S rRNA gene sequences through PCR from the gonads and spawned eggs of the shipworm *Bankia setacea*. The inability to detect symbionts outside of the host supports the vertical mode of transmission. On the other hand, genomic analysis of *T. turnerae* suggests a facultative symbiosis (Yang et al., 2009). More in depth localization studies, such as in situ hybridization or electron microscopy in conjunction with PCR-based detection of developing eggs/larvae are

needed. But a vertical mode of transmission suggests that segregation might occur during the early stages of development of the shipworm. Experimental colonization of aposymbiotic juveniles in the lucinid bivalve *Codakia orbicularis* might provide a possible mechanism of bacteriocyte colonization in teredinid bivalves. In the lucinid bivalve, symbionts enter into undifferentiated cells of gill filaments by endocytosis, which progressively differentiate into bacteriocytes (Gros et al., 1998). As juveniles develop, more bacteriocytes begin to appear. In the shipworm, this colonization might be mediated by metabolites produced by symbionts such as *T. turnerae* (Clade 1), creating a zone of inhibition, as seen in *in vitro* experiments, allowing itself exclusive access to developing bacteriocytes (Figure 4.3). Symbionts that are not able to effectively compete with Clade 1 symbionts, due to either iron restriction by Clade 1 siderophores or susceptibility to Clade 1 antimicrobials, are left to colonize developing bacteriocytes that are "unprotected" by Clade 1's zone of inhibition. Studies such as that of Gros et al. (1998) need to be conducted in teredinid bivalves to confirm a similar mechanism of bacteriocyte colonization.

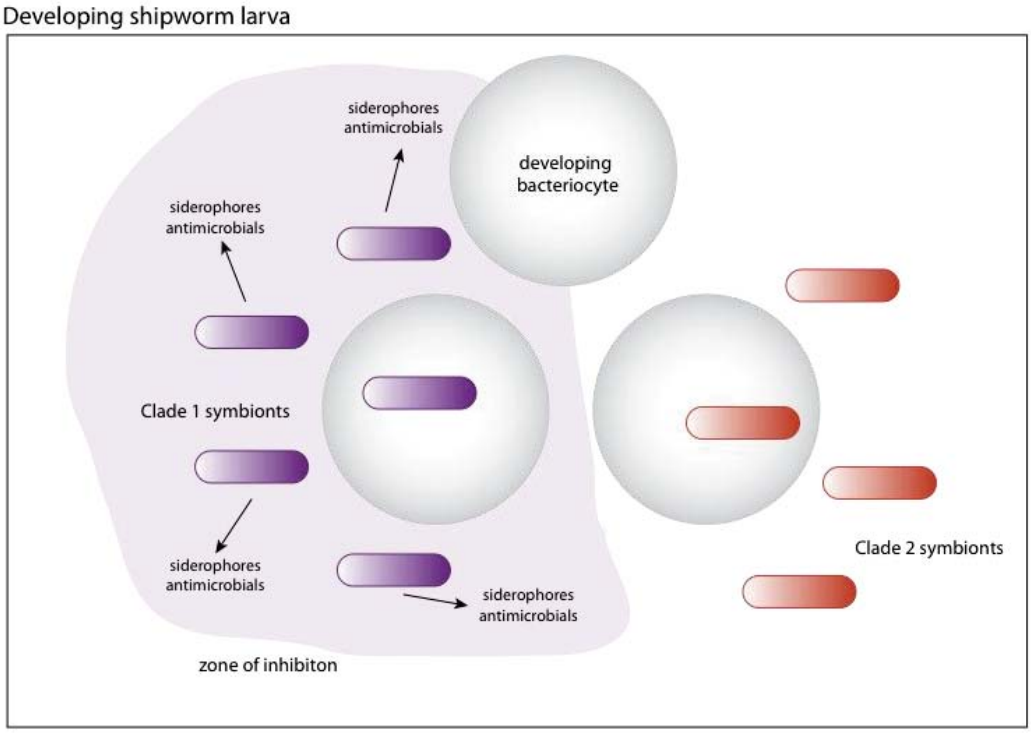

Figure 4.3. Proposed model of bacteriocyte colonization by symbionts.

Contained within bacteriocytes, symbionts are excluded from the surrounding environment by the bacteriocyte membrane. Transfer of metabolites and proteins seems to involve transport across these host membranes. Examination of aphid gene expression showed the upregulation of multiple transporters in bacteriocytes, suggesting these transporters are critical in metabolite exchange between the symbiont *Buchnera aphidicola* and the host (Hansen and Moran, 2011). Work with gills of the marine mussel *Mytilus californianus* show the use of a transporter to mediate transport of the amino acid taurine between intracellular compartments of gill cells and hemolymph (Neufeld and Wright, 1995). The packaging of symbiont metabolites or proteins might occur through a blebbing process, or outer membrane vesicle (OMV) formation, a process conserved among gram-negative bacteria (Kulp

and Kuehn, 2010). OMV's can differentially package select enzymes and receptors that contribute to nutrient acquisition, including iron-transport outer membrane proteins (Choi et al., 2011). Mashburn and Whitely (2005) showed that OMV's were critical to cell-to-cell communication and coordinated group behavior in the pathogen *Pseudomonas aeruginosa*. Coordinated behavior is particularly important with molecules that will diffuse away from cells, such as siderophores. Biosynthesis and excretion of siderophores would be most useful when there are enough cells to efficiently scavenge the metal (Keller and Surette, 2006). OMV's were visualized using electron microscopy by bacteriocyte-enclosed symbionts of the cockroach *Cryptocercus punctulatus* (Bigliardi et al., 1995). OMV's may play a critical role in the shipworm symbiont community as well. Delivery of OMV's to other bacteriocytes or distant parts of the shipworm might utilize the circulatory system, excreting OMV's into the hemolymph of the proximal branchial vein (refer to Figure 1.2). Targeted delivery of OMV's is possible through binding to host receptors (Kulp and Kuehn, 2010) and the mechanism of delivery has been shown to occur by endocytosis (Kesty et al., 2004).

The findings of this work suggest that universal cooperation between endosymbionts of shipworms does not occur. Cross-streaking experiments show the inhibition of multiple clades of symbionts by symbiont Clade 1, most likely the result of multiple mechanisms. Disc-diffusion assays show that siderophores alone can inhibit the growth of other symbionts. These results were used to propose a possible mechanism of segregation found in the gills of the host. However, these experiments also revealed the presence of social cheaters in the community. Bs02 is able to utilize

the heterologous siderophore, triturnerbactin, while not producing a siderophore of its own. This suggests that some symbiont members do not contribute equally to the good of the community.

5 Conclusions

Shipworms harbor a closely related group of endosymbionts that are believed to supplement the host's nutrition. The symbionts produce cellulolytic enzymes that aid in the digestion of wood and supplement this carbon-rich diet of wood by providing a source of nitrogen through nitrogen fixation. The burden of acquiring iron within the host may be exacerbated by the additional iron requirements of nitrogen fixation.

A strategy used by many microorganisms in the acquisition of iron is the production of siderophores. The recent genome sequencing of the shipworm symbiont *T. turnerae* T7901 revealed an extensive secondary metabolome that includes a putative siderophore biosynthetic gene cluster. This work examined this siderophore biosynthetic gene cluster and the resulting siderophore product. The chemical isolation of the siderophore revealed a novel structure of 2,3-DHB-L-ornithyl-L-serine, which is closely related to the previously described siderophores vanchrobactin and chrysobactin. Together these siderophores form a catecholate class of siderophores with a DHB binding group, a hydrophilic, positively charged amino acid spacer, and serine backbone. The structure of the turnerbactin biosynthetic gene cluster showed what appeared to be a relatively straightforward biosynthesis scheme. But upon further examination, the catalytic domains of the NRPS are not completely amenable to the current bioinformatic predictive capabilities for NRPS's. The addition of the analyses on turnerbactin biosynthesis will aid future research in the predictive analyses of NRPS domains. NRPS's have

the potential to produce a wide array of medically relevant and therapeutic compounds. The search for novel compounds is ongoing, and increasingly accurate predictive capabilities allow for more efficient genomic-based screening of pathways that might yield novel-acting compounds.

After the characterization of *T. turnerae* T7901's siderophore, a siderophore profile for the shipworm symbiont community was created. Multiple strains of *T. turnerae* have been discovered from numerous genera and species of shipworms. A survey of *T. turnerae* strains was carried out to determine turnerbactin production by this species of shipworm symbiont. Turnerbactin was found in all strains of *T. turnerae* examined. Additionally, the presence of turnerbactin was detected in *Lyrodus pedicellatus*, a shipworm known to contain *T. turnereae*. Combined, these results suggest that turnerbactin might be found in numerous shipworm species and may play a role in the symbiosis. Other symbiont groups were also examined for siderophores through genomic analysis. These findings suggest a theme of catecholate siderophore amongst shipworm symbionts.

Finally, the potential for siderophores to affect community dynamics within the symbiotic community was presented. These experiments show that siderophores can have antagonistic actions towards other symbionts, but that they are most likely part of a multiple-faceted mechanism of antagonism. A model of segregation within the gill by different symbiont groups was proposed using these results. Also discovered in these experiments was the presence of a social cheater in the community. These symbionts seem to be benefitting from the other members of the community without expending comparable resources.

Future directions

This work provided an overview of the siderophore potential of shipworm symbionts. The role of siderophores in the interactions between symbionts was presented. In addition to symbiont interactions, siderophores may also play other key roles in the shipworm system. The digestive organs of the shipworms are relatively devoid of microbes (Greenfield and Lane, 1953; Betcher, 2011), which is in contrast to other wood-digesting organisms. It is believed that symbiont-derived compounds are transported into the digestive organs to aid in digestion. Siderophores may also be exported to these areas and play a role in nutrient deprivation to prevent unwanted microbes from colonizing these organs.

Much is unknown about the exchange of metabolites and proteins in the shipworm system, both symbiont-host exchange as well as symbiont-symbiont exchange. Localization studies such as the multiple-isotope imaging mass spectrometry study of nitrogen fixation in the gills of *Lyrodus pedicellatus* (Lechene et al., 2007) are valuable in determining the location and transport of symbiont-derived substrates. Additional MS-imaging studies in the shipworm system will reveal the fate of symbiont-derived compounds, such as siderophores, and lend insight into the possible roles for these compounds.

While the siderophore biosynthetic gene cluster was examined in this work, the genome of *T. turnerae* T7901 codes for eight other putative secondary metabolite gene clusters, which include both polyketide synthases and NRPS's. These

compounds may play an integral role in the symbiosis and might also provide interesting biological activity for biomedical applications. The genetic techniques developed for *T. turnerae* T7901 in this work will prove valuable in assigning these additional gene clusters to their isolated products.

Introduction of mutants into the shipworm system would provide valuable information. For example, the effects of colonization and persistence by a siderophore-production mutant could be monitored and analyzed to determine how the production of this compound affects the relationship between symbiont and host. However, the mechanism of transmission of symbionts has not yet been firmly established, which leaves us uncertain as to how to introduce a strain of choice. In addition, the curing of shipworms of any bacterial symbionts has never been carried out, leaving any results of successful introduction of strains inconclusive, as natural strains would still be present in the shipworm.

Transmission of the endosymbiontic community in teredinid bivalves requires further study. Sipe et al. (2000) propose a vertical mode of transmission, but genomic analysis of *T. turnerae* T7901 also suggests the potential for horizontal transmission (Yang et al., 2009). A more in depth study involving microscopy along with PCR-based detection would provide further evidence. Heavy focus should be on the developmental stages of the shipworm, capturing the formation of bacteriocytes within the gill. Once that is established, focus can be shifted toward the development of bacteriocytes during the shipworm's maturation process.

Through the current work, the mechanism of segregation of symbionts within the gill is hypothesized to involve secondary metabolites produced by shipworm

symbionts. The potential for such segregation was shown *in vitro*, but more studies are needed to determine the exact mechanism of symbiont segregation in the gills. Gene expression analysis of axenic cultures of *T. turnerae* has been examined (Fishman, 2010). These studies can be extended to include gene expression analysis of symbiont strains in co-culture with other symbionts to determine if expression of secondary metabolite gene clusters is upregulated in the presence of other symbionts.

The catecholate-structural features of turnerbactin suggest that turnerbactin is able to bind other metals aside from iron, such as molybdenum (Duhme-Klair, 2009). Similar catecholate siderophores were shown to scavenge molybdenum for the soil diazotroph, *Azotobacter vinelandii*, and regulate toxicity in the presence of tungsten (Bellenger et al., 2008; Wichard et al., 2008). Molybdenum is another cofactor of the enzyme nitrogenase and similar roles for turnerbactin may occur with *T. turnerae* during times of nitrogen fixation.

Uptake experiments with turnerbactins and other symbiont siderophores would also be useful. The results of experiments such as radiolabeled uptake experiments would provide compelling evidence for the uptake of exogenous siderophores by gill symbionts, as well as the uptake of siderophore-bound trace metals, such as iron and molybdenum.

Appendix A. Supplementary NMR figures

Table A.1. NMR data for Diturnerbactin (**2**) and Dehydrated diturnerbactin (**4**) (800 MHz) in CD_3OD.

	Diturnerbactin (2)			Dehydrated diturnerbactin (4)			
Position	**δ_H (*J* in Hz)**	**HSQC**	**HMBC**	**δ_C**	**δ_H (*J* in Hz)**	**TOCSY**	**HMBC**
DHB							
1, 1'				170.78 170.88			
2, 2'				117.21 117.06			
3, 3'				149.49 149.37			
4, 4'				147.26 147.17			
5, 5'	6.98, dd (1.6, 2.4), [1H] 6.97, dd (1.6, 2.4), [1H]	118.6	2, 3, 4, 6, 7	119.93 119.86	6.99, dd (8.0, 1.6), [1H] 6.95, dd (8.0, 1.6), [1H]	6, 7	2, 3, 4, 6 2', 3', 4', 6'
6, 6'	6.77, t (8.0) [2H]	118.8	1, 2, 3, 4, 5, 7	120.00 119.89	6.78, t (8.0), [2H] 6.72, t (8.0), [2H]	5, 7	2, 3, 4, 5, 7
7, 7'	7.38, dd (1.6, 8.0), [1H] 7.36, dd (1.6, 8.0), [1H]	118.4	1, 3, 4, 5, 6	119.71 119.66	7.37, dd (8.0, 1.6), [1H] 7.31, dd (8.0, 0.8), [1H]	5, 6	1, 3, 4, 6 1', 3', 4', 6'
Ornithine							
8, 8'				173.67 172.27			
9	4.79, m [1H]	51.93	1, 8, 10, 11	54.37	4.82, m, [1H]	10, 11, 12	1, 8, 10, 11
9'	4.76, m, [1H]	52.53	1', 8', 10', 11'	53.94	4.73, m, [1H]	10', 11', 12'	1', 8', 10', 11'
10, 10'	2.05, m, [2H] 1.88, m, [2H]	28.8	8, 9, 11, 12	29.89, 29.78	2.07, m, [2H] 1.92, m, [2H]	9, 11, 12 9, 11, 12	8, 9, 11, 12 8, 9, 11, 12
11, 11'	1.81, m [4H]	23.5	9, 10, 12	25.03, 24.97	1.82, m, [4H]	9, 10, 12	9, 10, 12
12, 12'	2.99, m [4H]	38.9	10, 11	40.30, 40.27	3.00, m, [4H]	9, 10, 11	10, 11
Serine							
13 13'				164.44 172.23			
14	4.58, t (4.0, 4.8), [1H]	54.96	8, 13, 15	133.02	4.88, m, [1H]		8', 13', 15'
14'	4.80, m, [1H]	51.97	8', 13', 15'	53.12			
15	3.95, dd (4.8, 12.0), [1H] 3.82, dd (4.8, 11.2), [1H]	61.23	13, 14 13, 14	112.57	6.23, s, [1H] 5.90, s, [1H]	15 15	13, 14 13, 14
15'	4.75, m, [1H] 4.43, dd (4.8, 11.2), [1H]	64.04	13, 13', 14' 13, 13', 14'	66.09	4.68, dd (11.2, 6.4), [1H] 4.53, dd (11.2, 6.4), [1H]	15' 15'	13, 13', 14' 13, 13', 14'

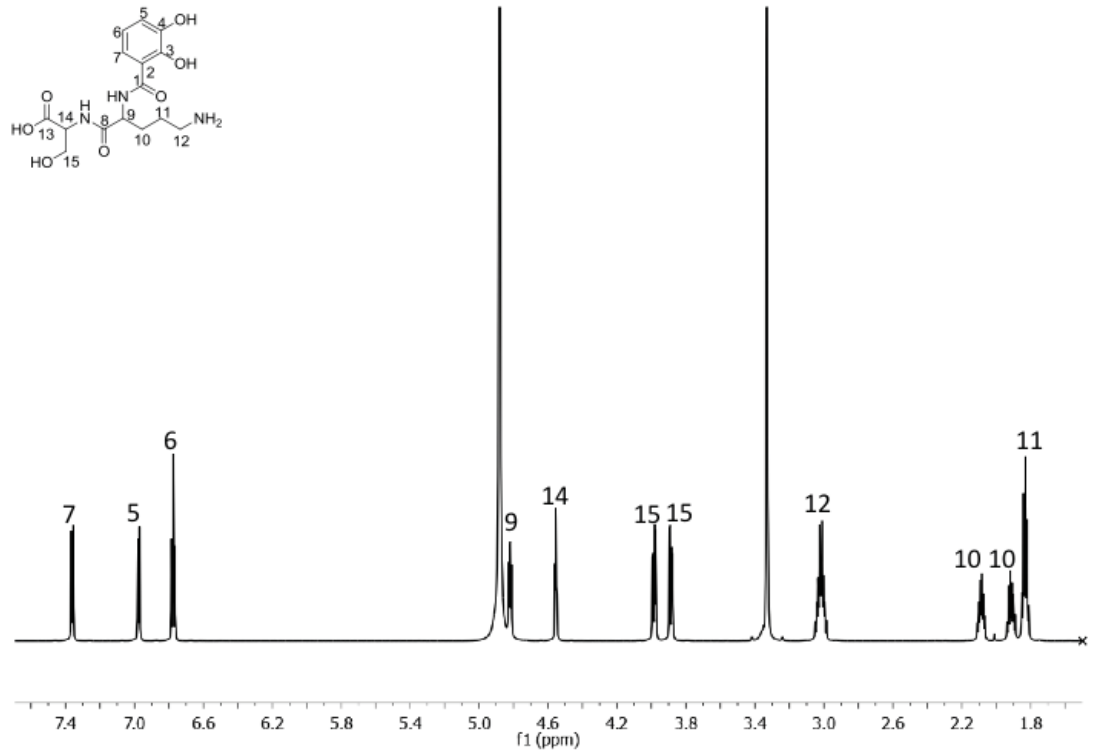

Figure A.1. Turnerbactin (**1**) ^{1}H NMR spectrum (800 MHz) in CD_3OD.

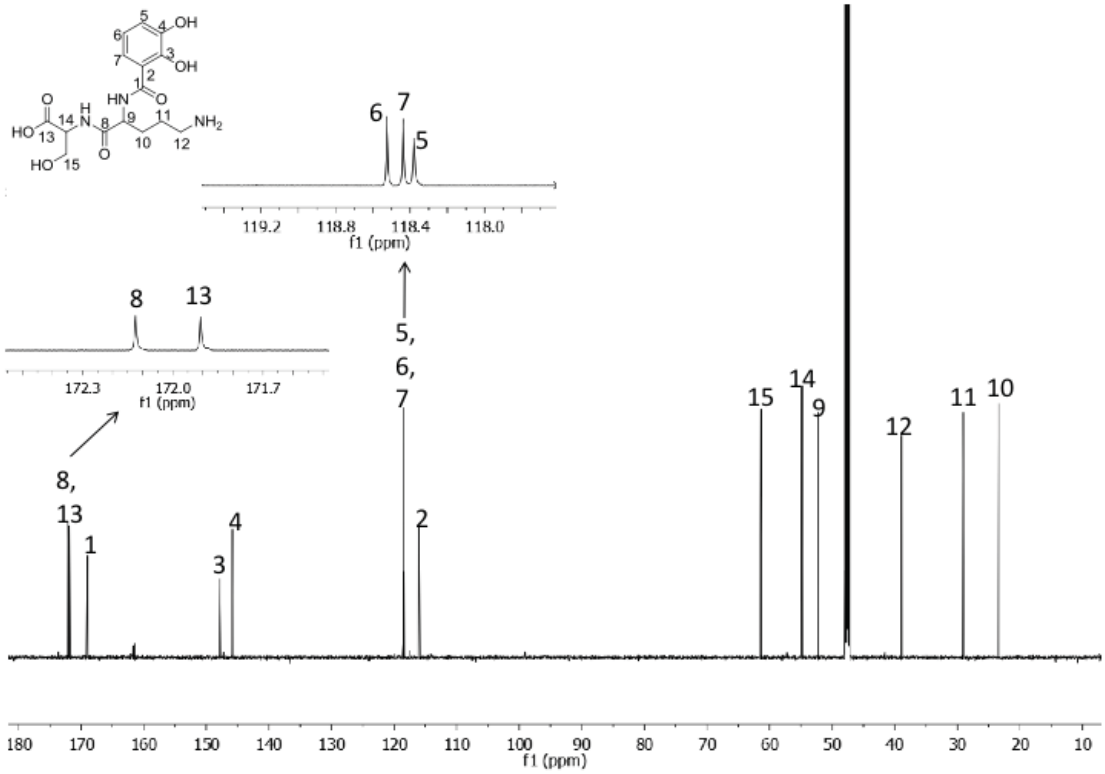

Figure A.2. Turnerbactin (**1**) ^{13}C NMR spectrum (800 MHz) in CD_3OD.

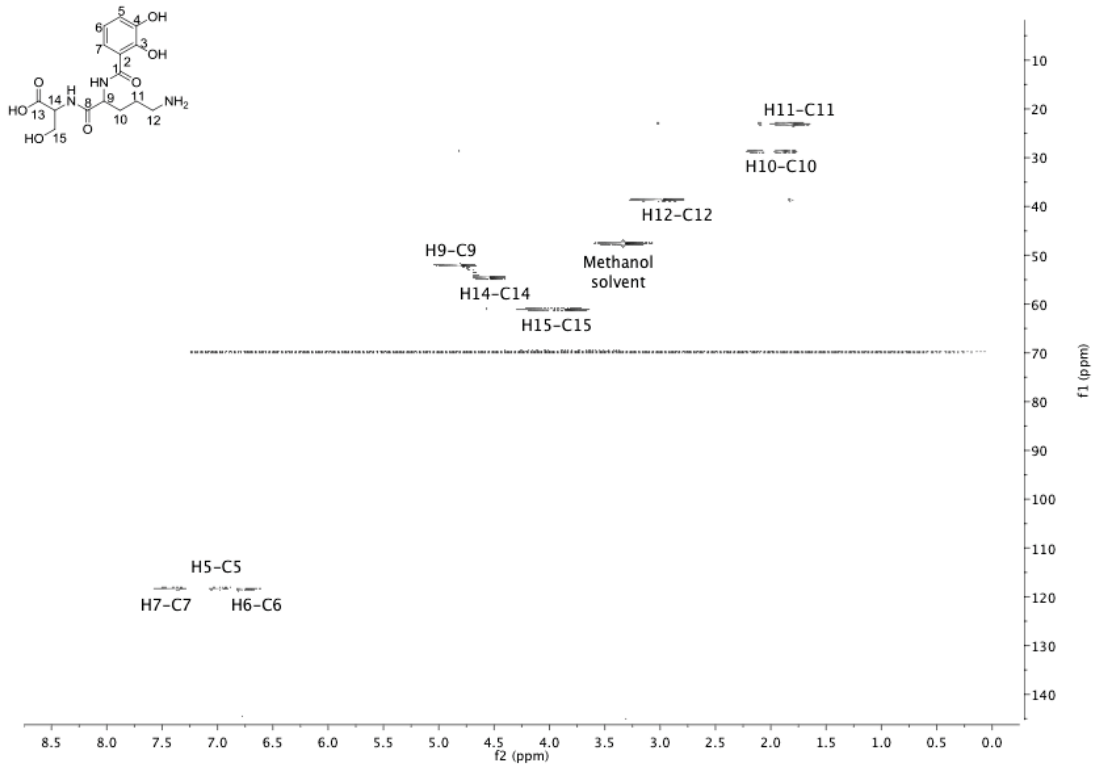

Figure A.3. Turnerbactin (**1**) ^{1}H-^{13}C HSQC spectrum (800 MHz) in CD_3OD.

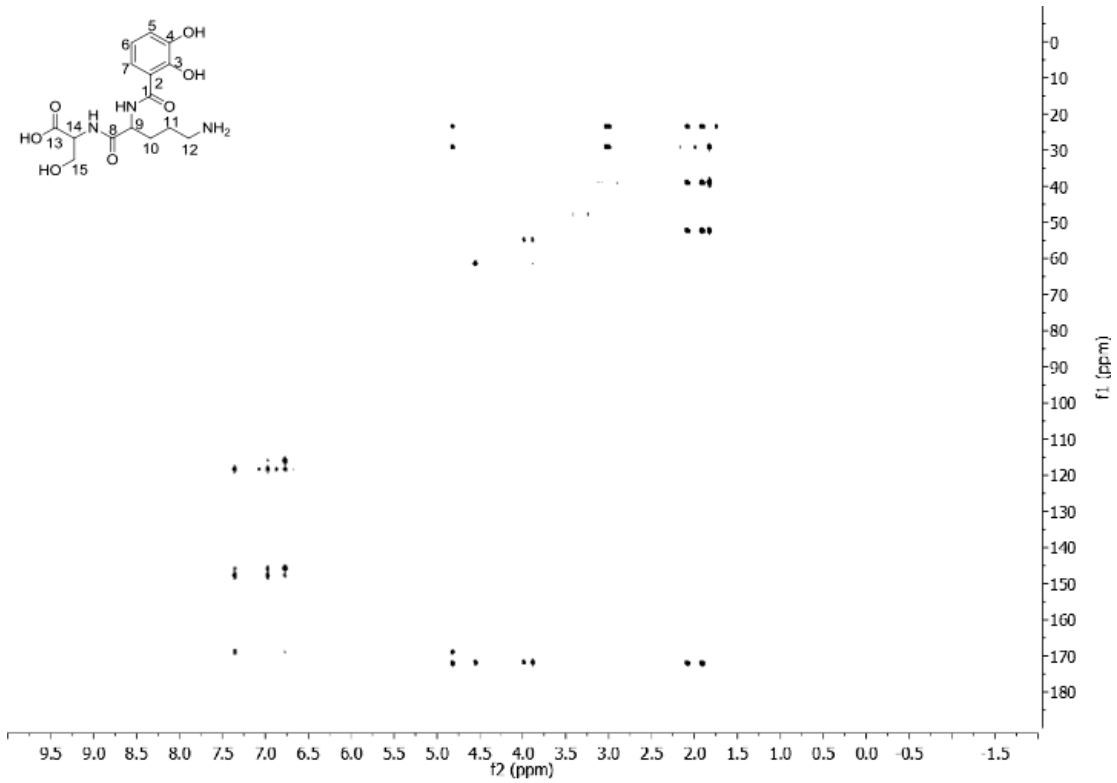

Figure A.4. Turnerbactin (**1**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD.

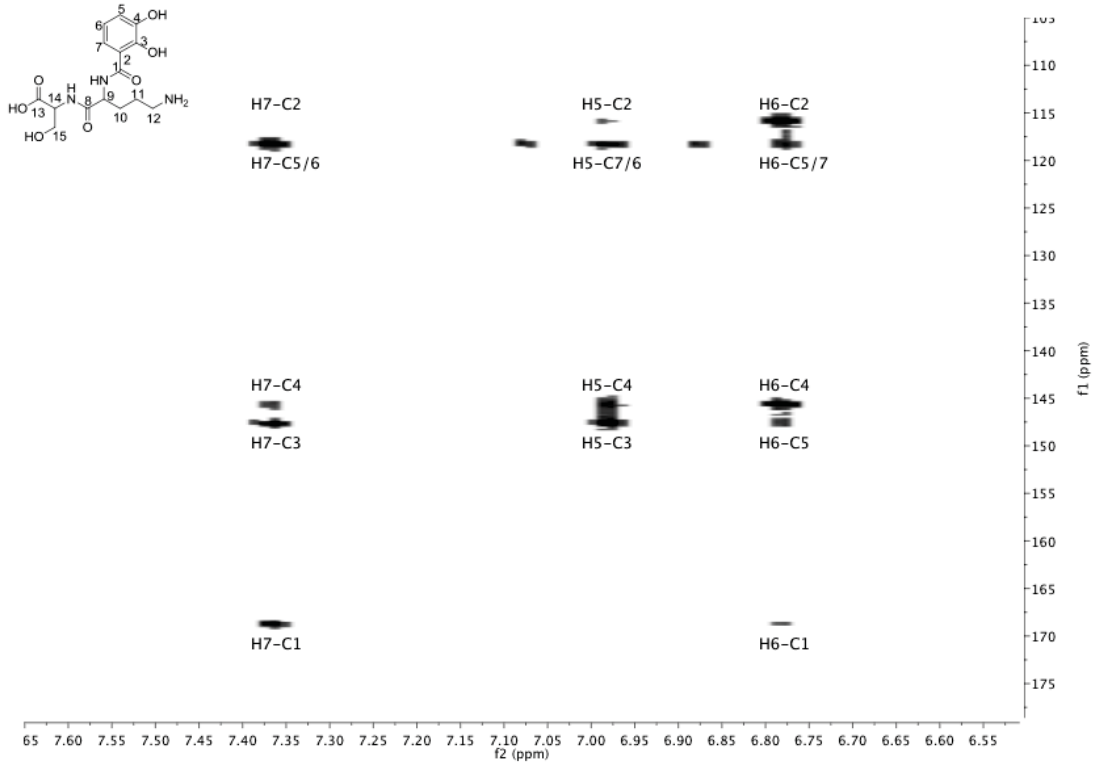

Figure A.5. Turnerbactin (**1**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

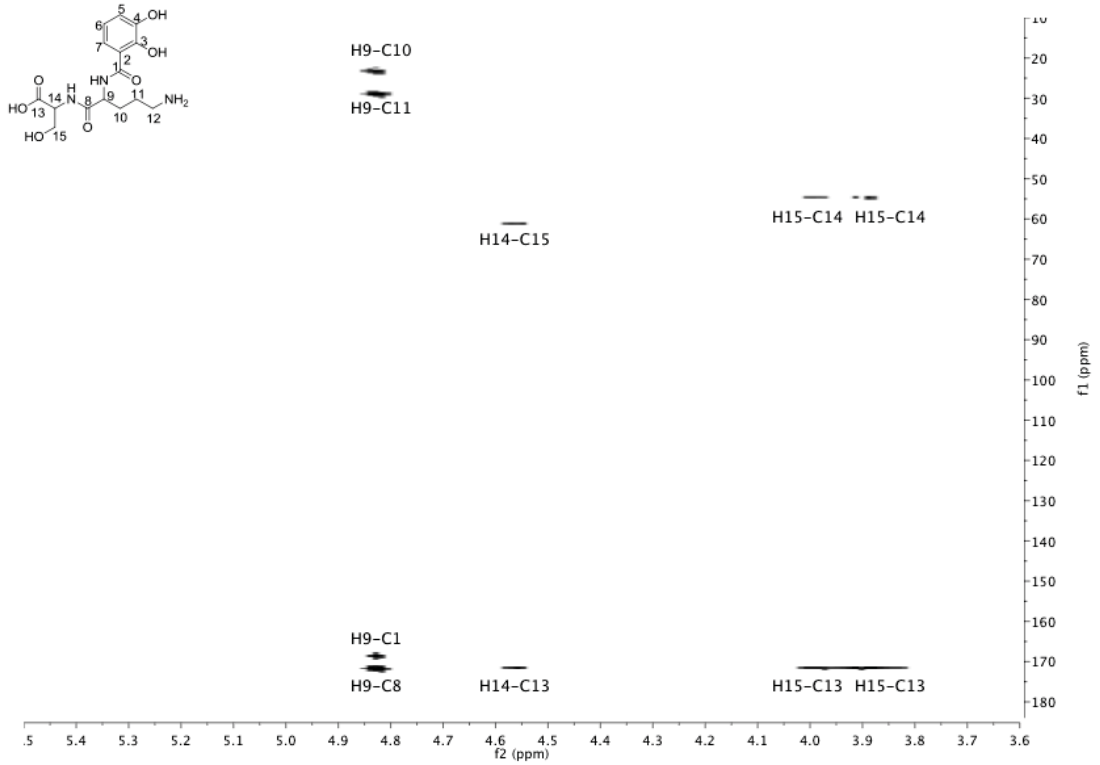

Figure A.6. Turnerbactin (**1**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

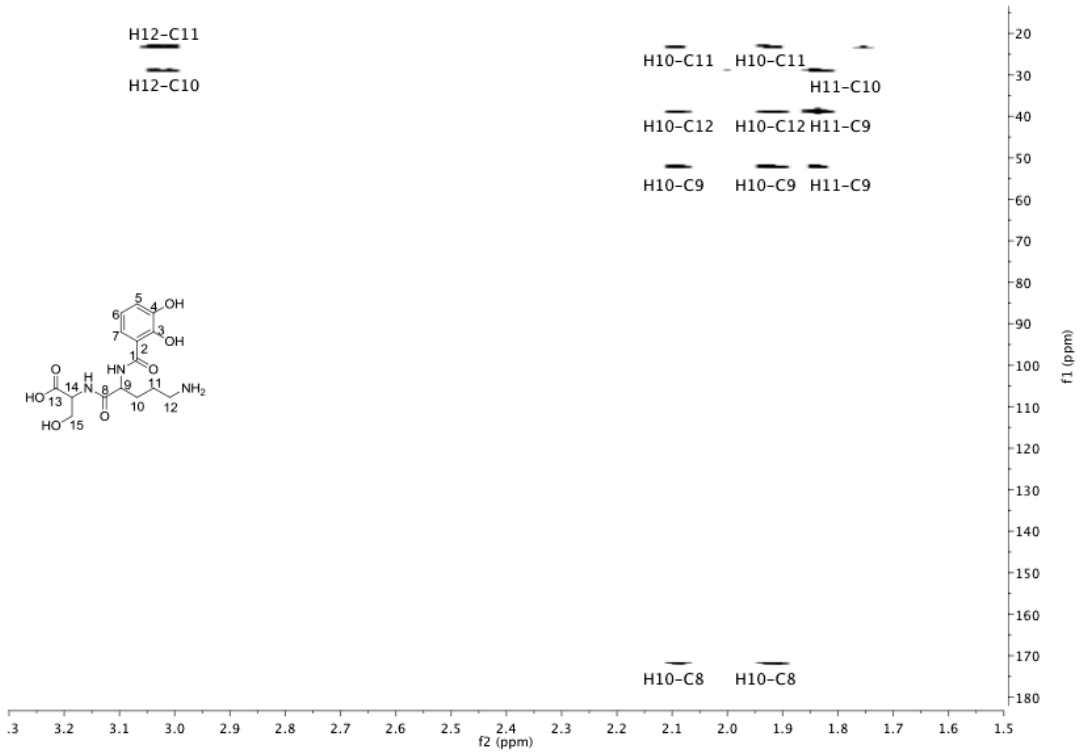

Figure A.7. Turnerbactin (**1**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

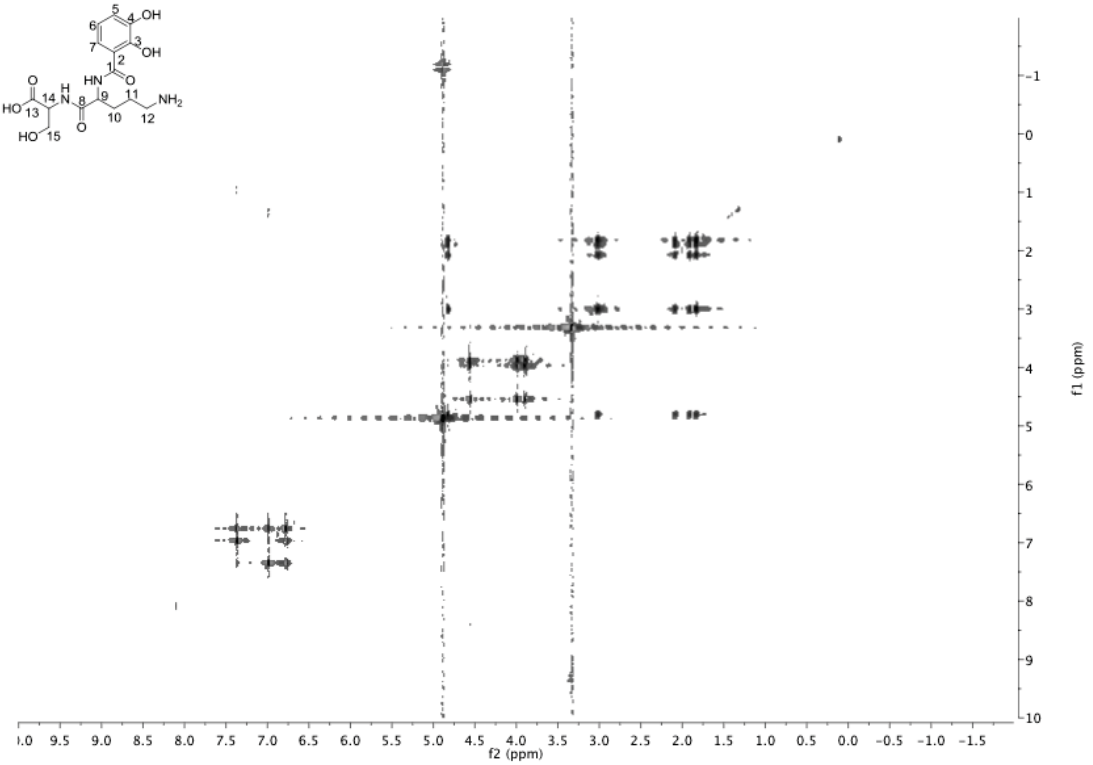

Figure A.8. Turnerbactin (**1**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD.

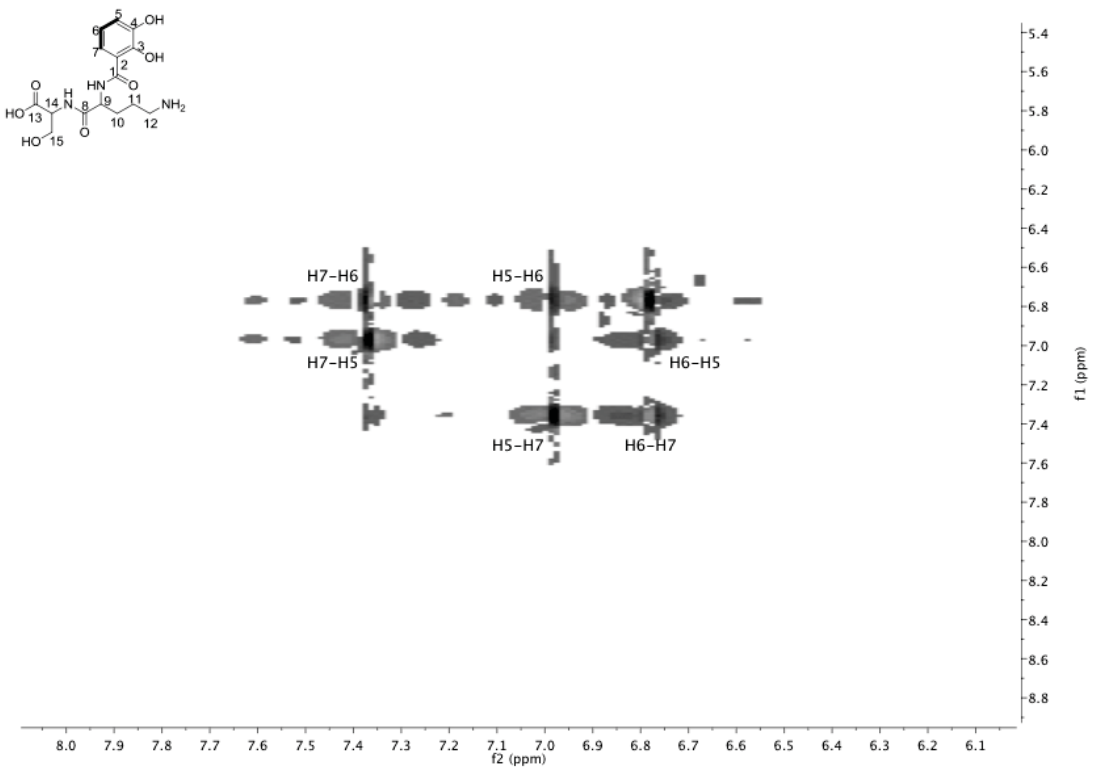

Figure A.9. Turnerbactin (**1**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

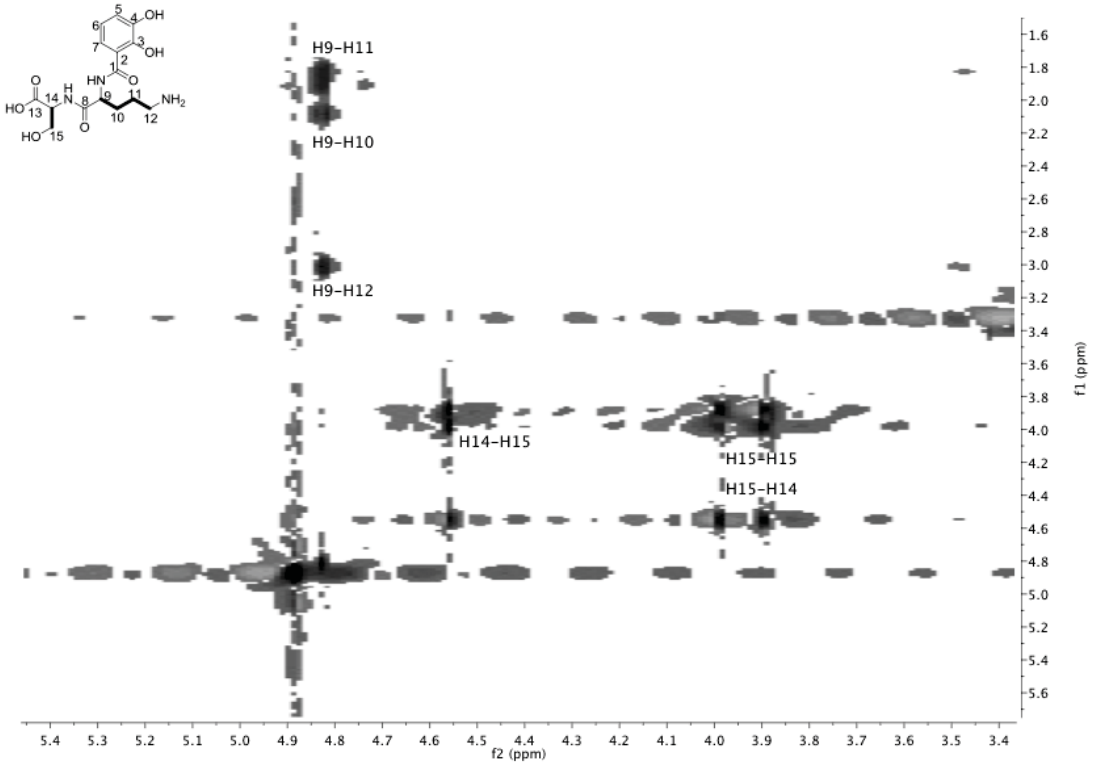

Figure A.10. Turnerbactin (**1**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

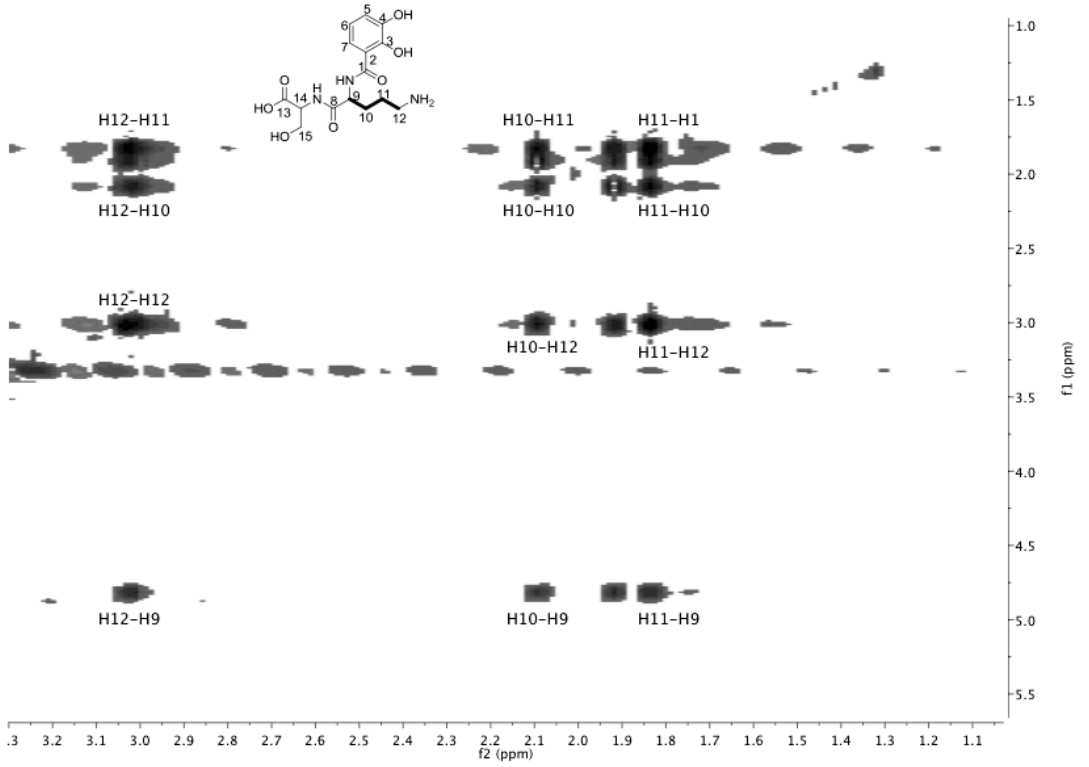

Figure A.11. Turnerbactin (**1**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

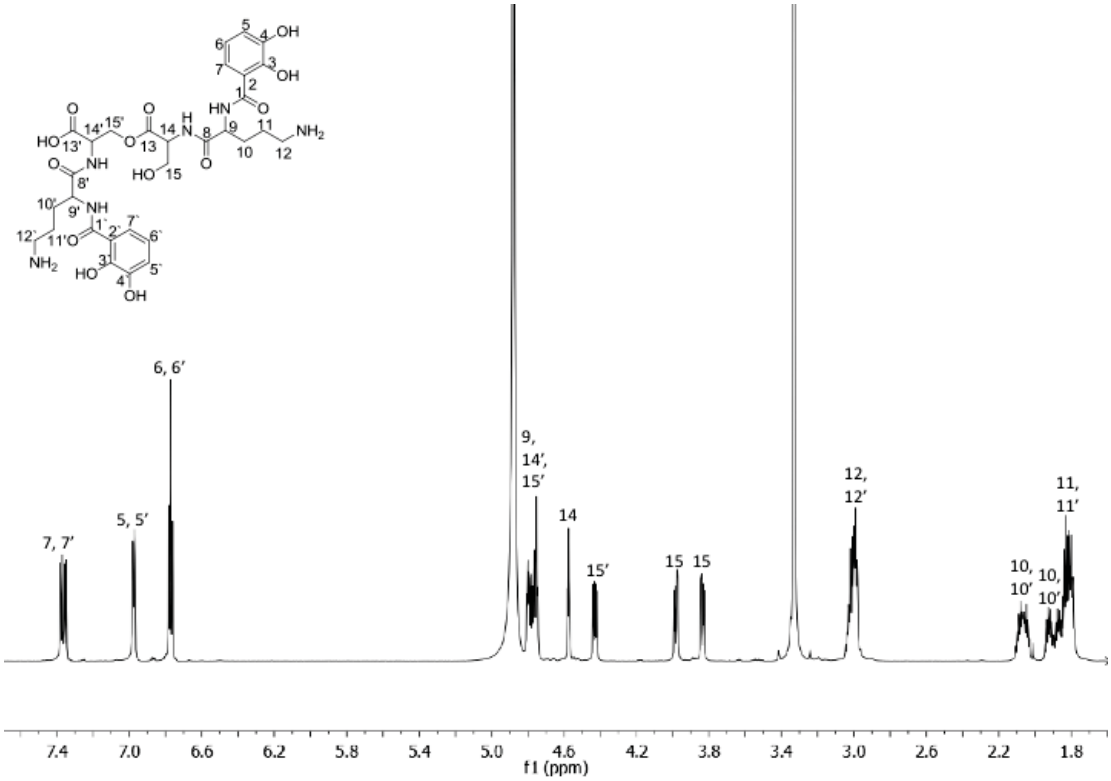

Figure A.12. Diturnerbactin (**2**) ^{1}H NMR spectrum (800 MHz) in CD_3OD.

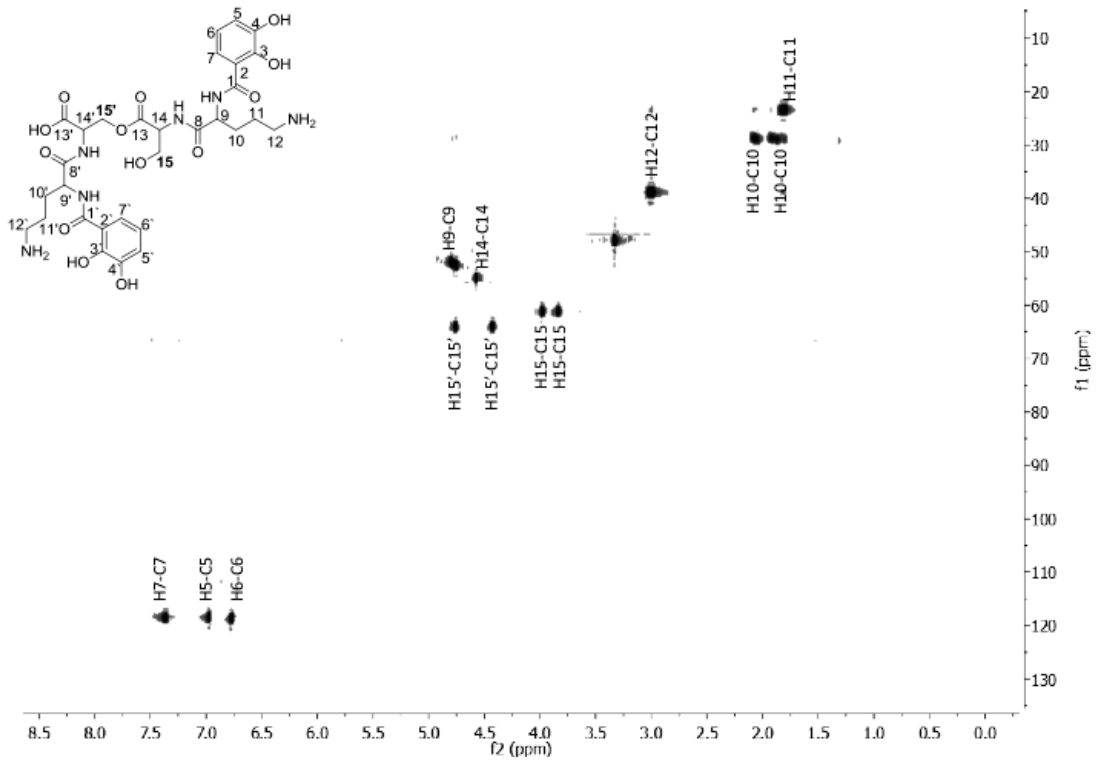

Figure A.13. Diturnerbactin (**2**) ^{1}H-^{13}C HSQC spectrum (800 MHz) in CD_3OD.

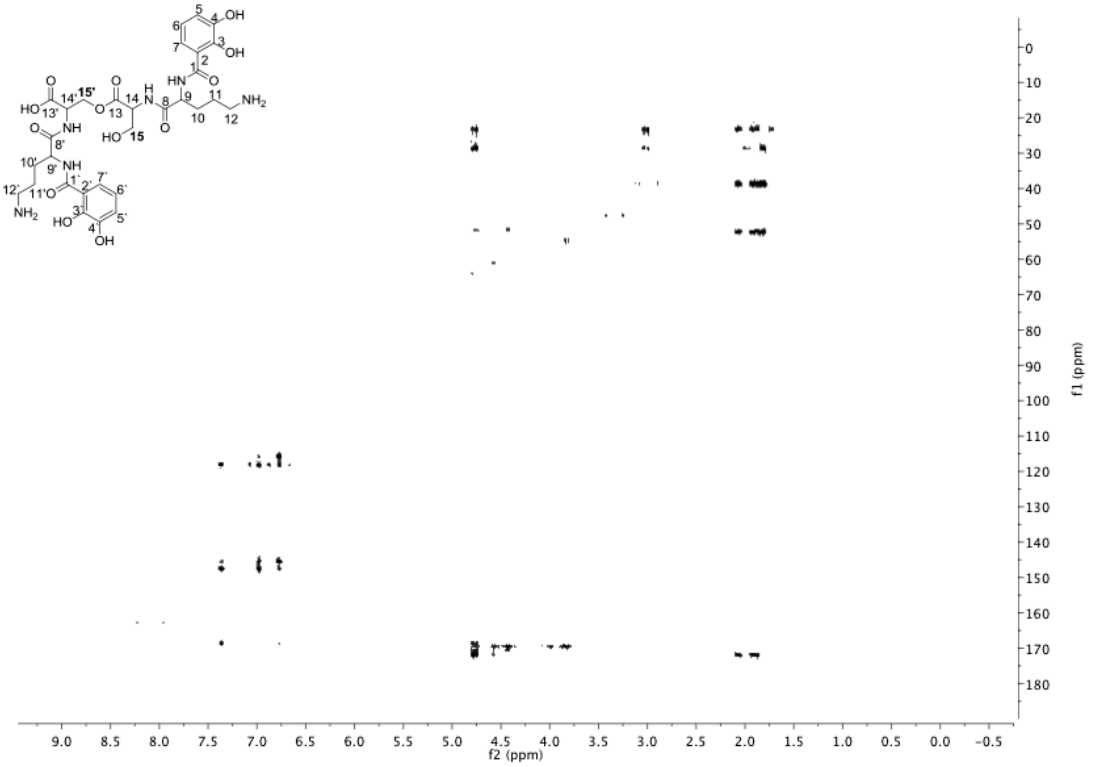

Figure A.14. Diturnerbactin (**2**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD.

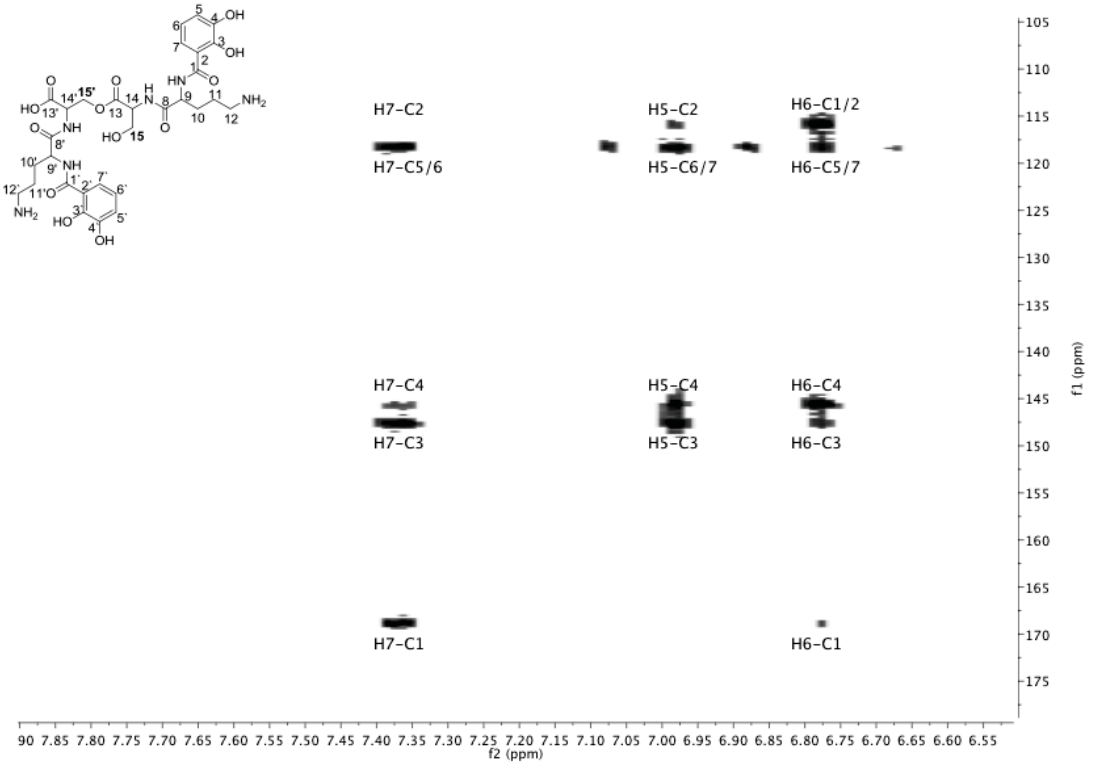

Figure A.15. Diturnerbactin (**2**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

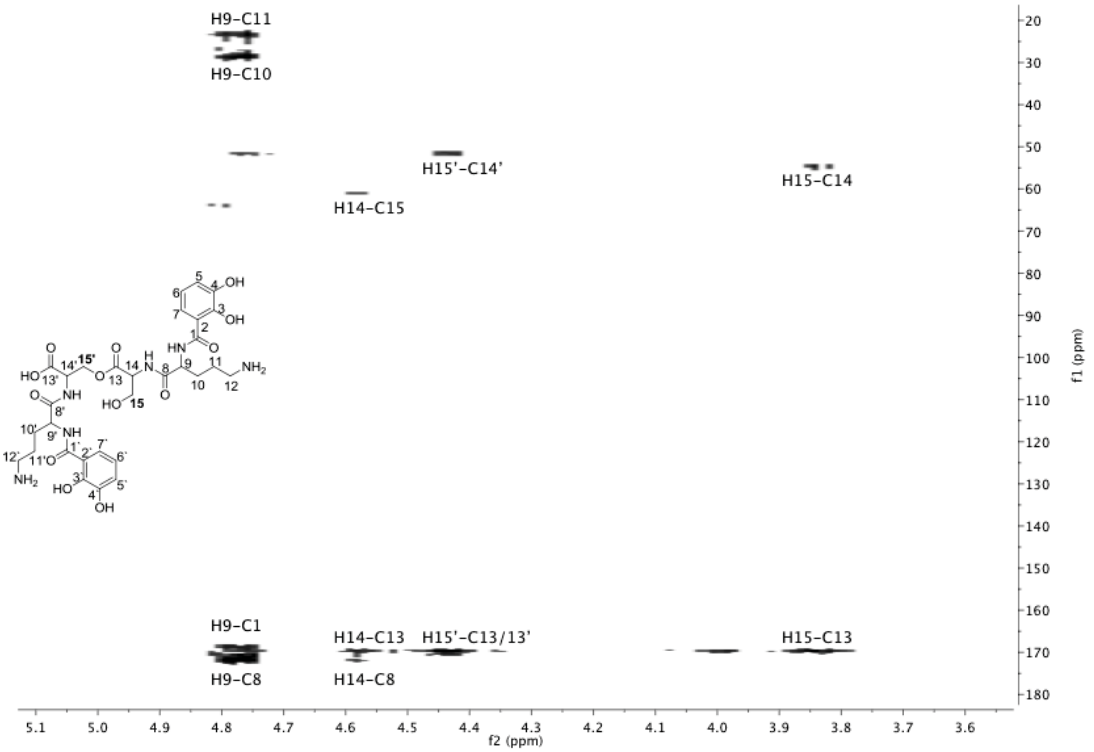

Figure A.16. Diturnerbactin (**2**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

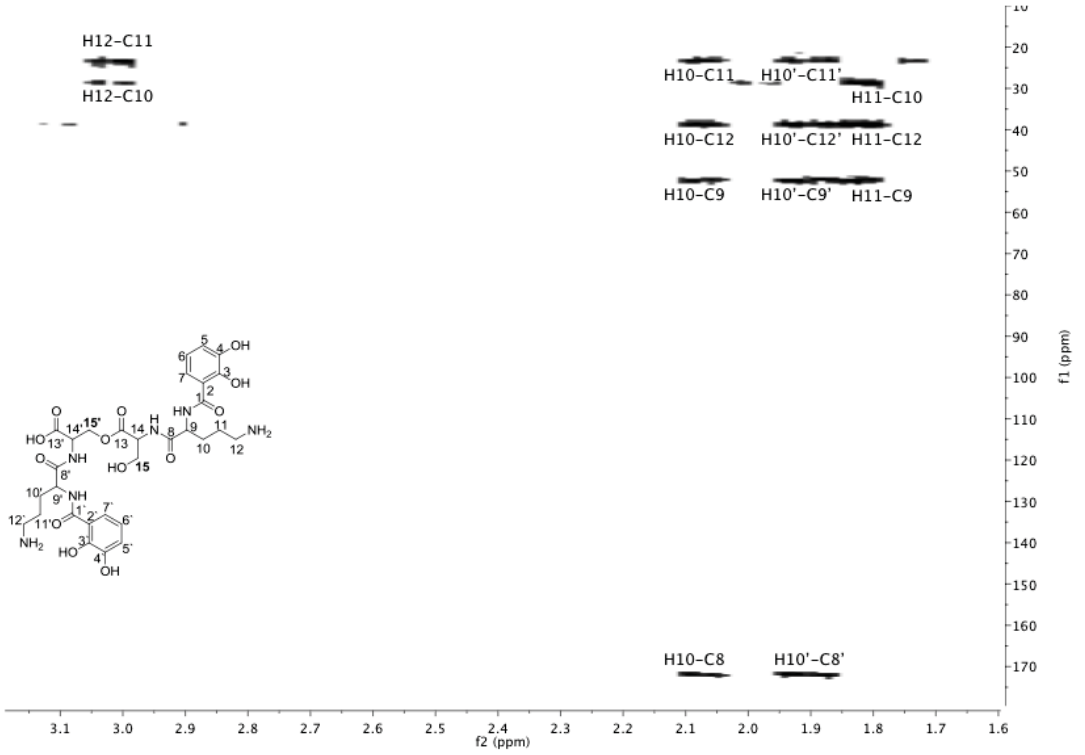

Figure A.17. Diturnerbactin (**2**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

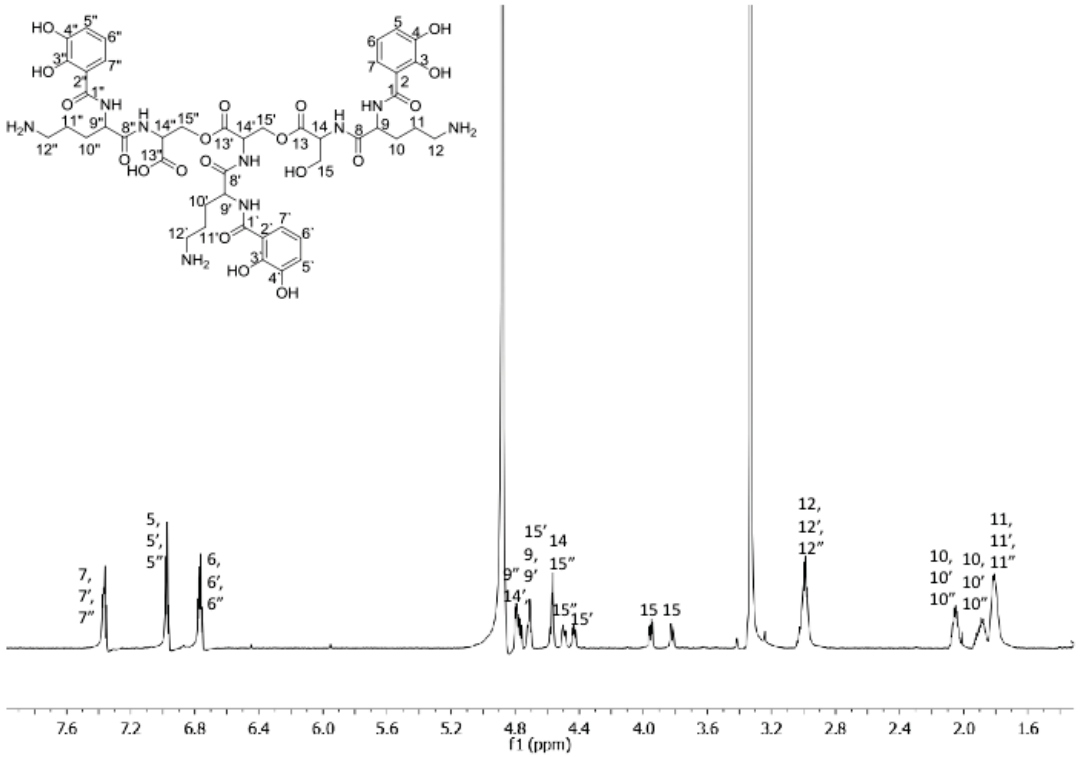

Figure A.18. Triturnerbactin (**3**) ^{1}H NMR spectrum (800 MHz) in CD_3OD.

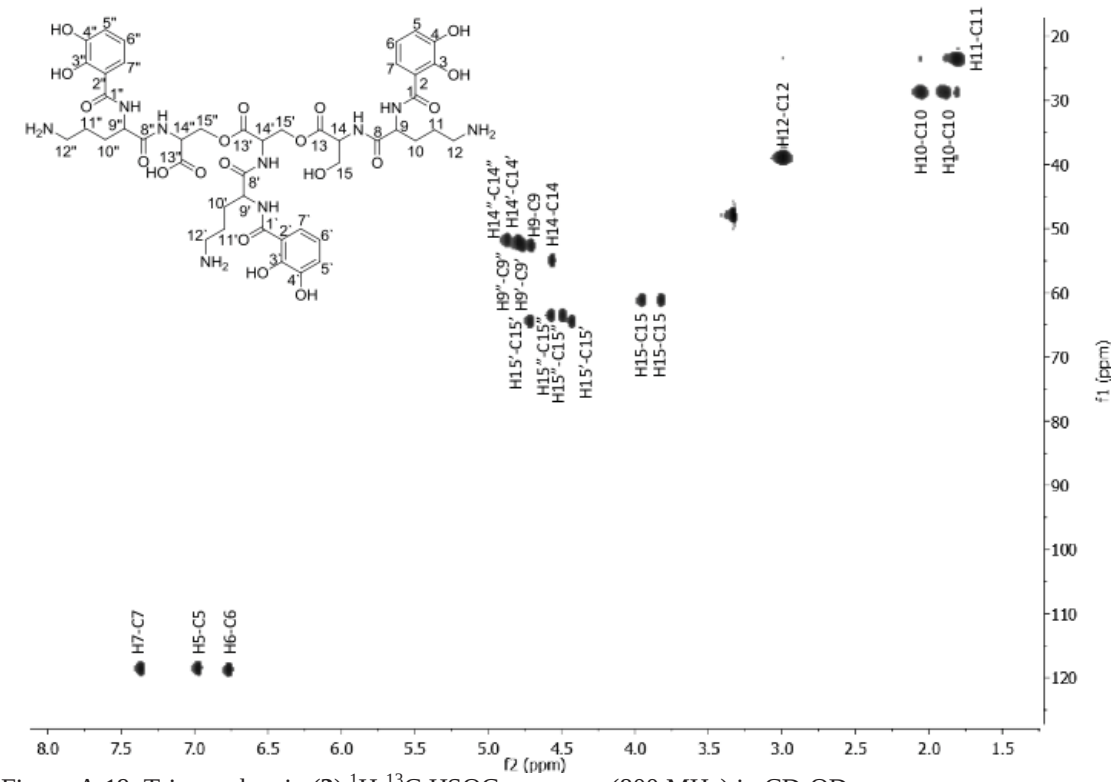

Figure A.19. Triturnerbactin (**3**) ^{1}H-^{13}C HSQC spectrum (800 MHz) in CD_3OD.

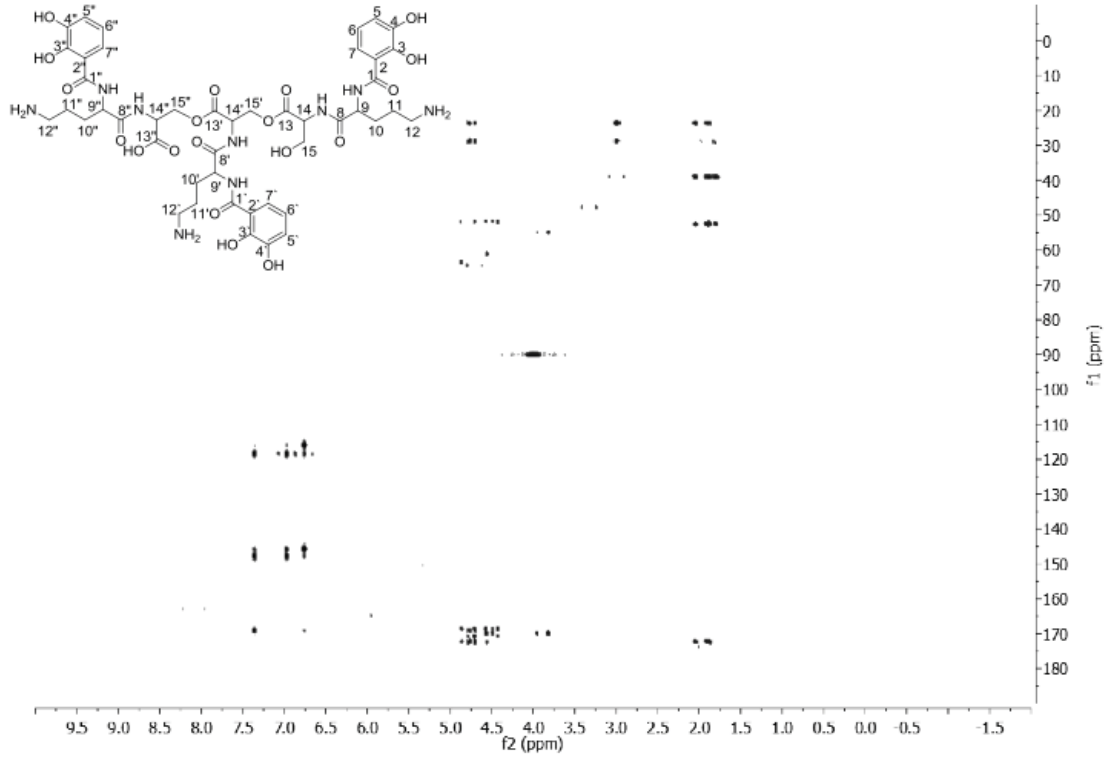

Figure A.20. Triturnerbactin (**3**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD.

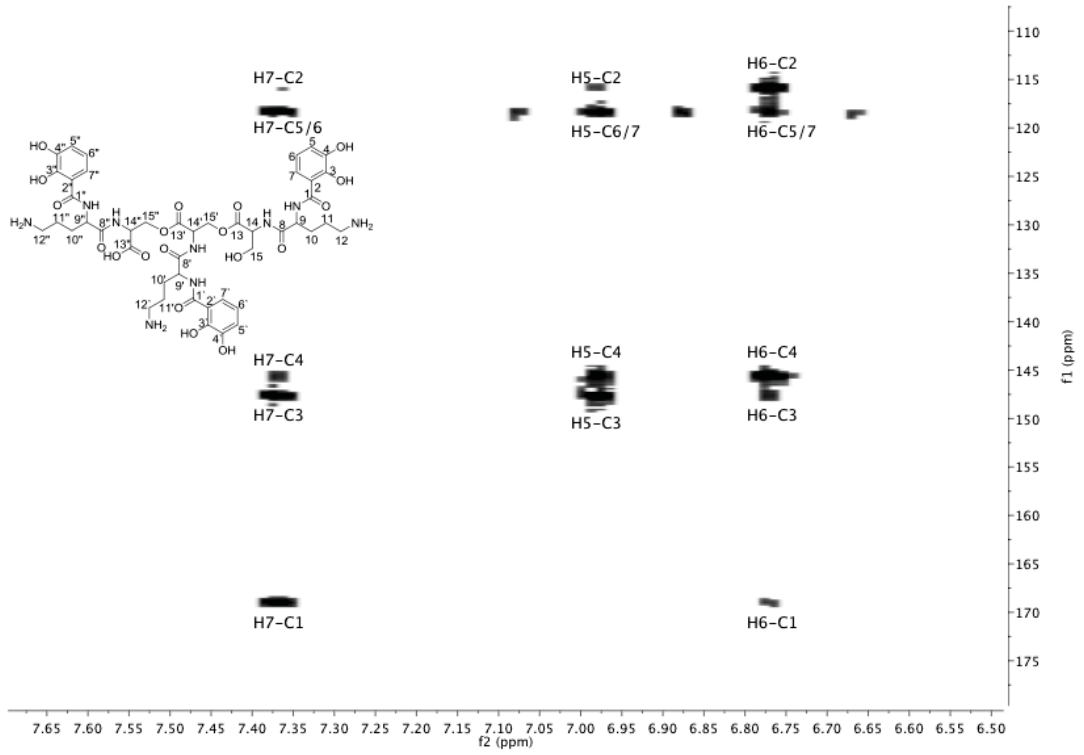

Figure A.21. Triturnerbactin (**3**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

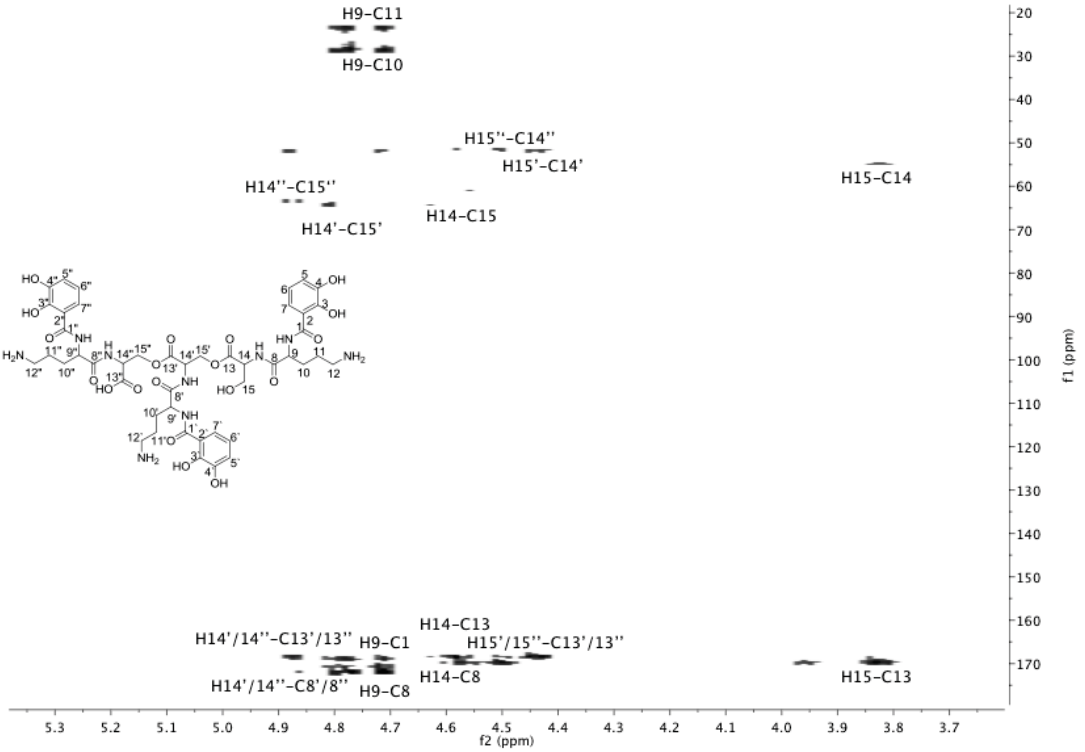

Figure A.22. Triturnerbactin (**3**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

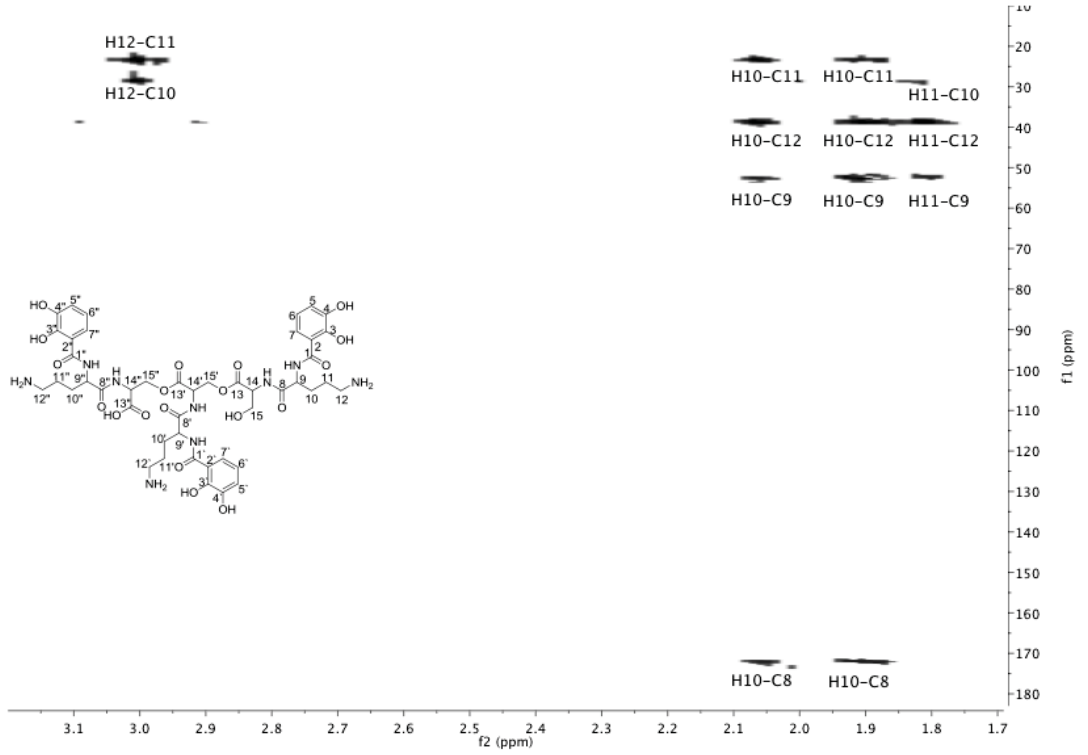

Figure A.23. Triturnerbactin (**3**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

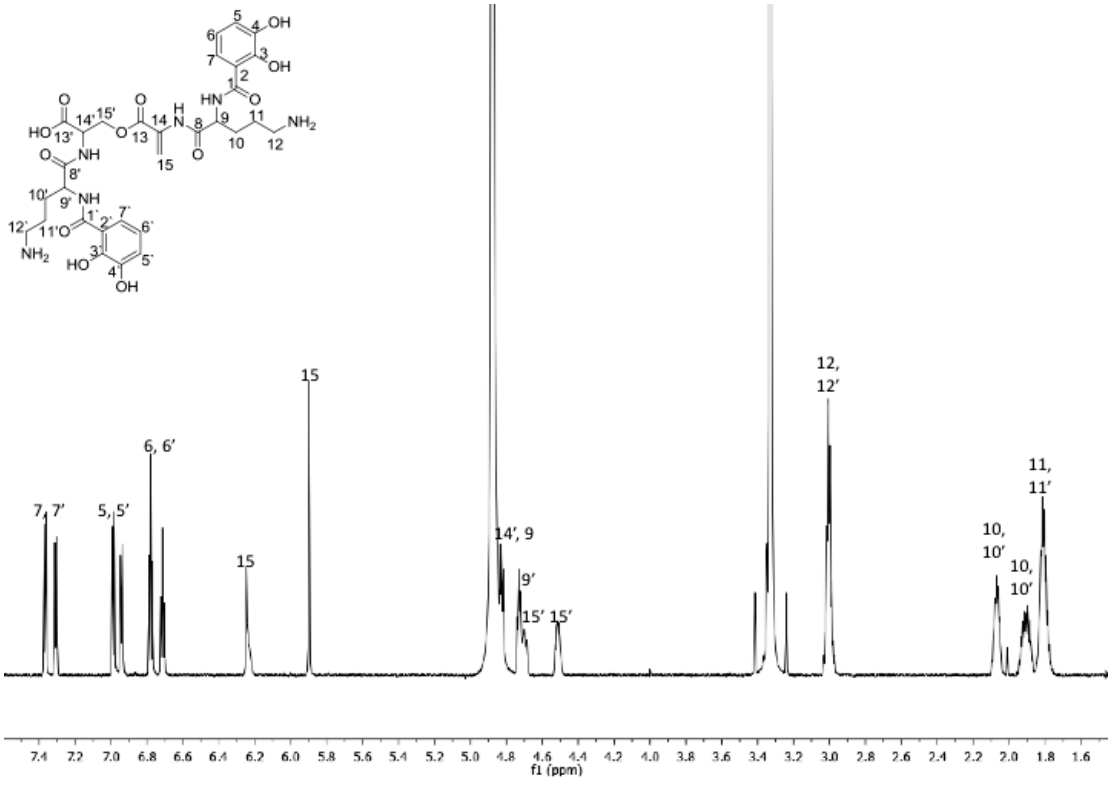

Figure A.24. Dehydrated diturnerbactin (**4**) ^{1}H NMR spectrum (800 MHz) in CD_3OD.

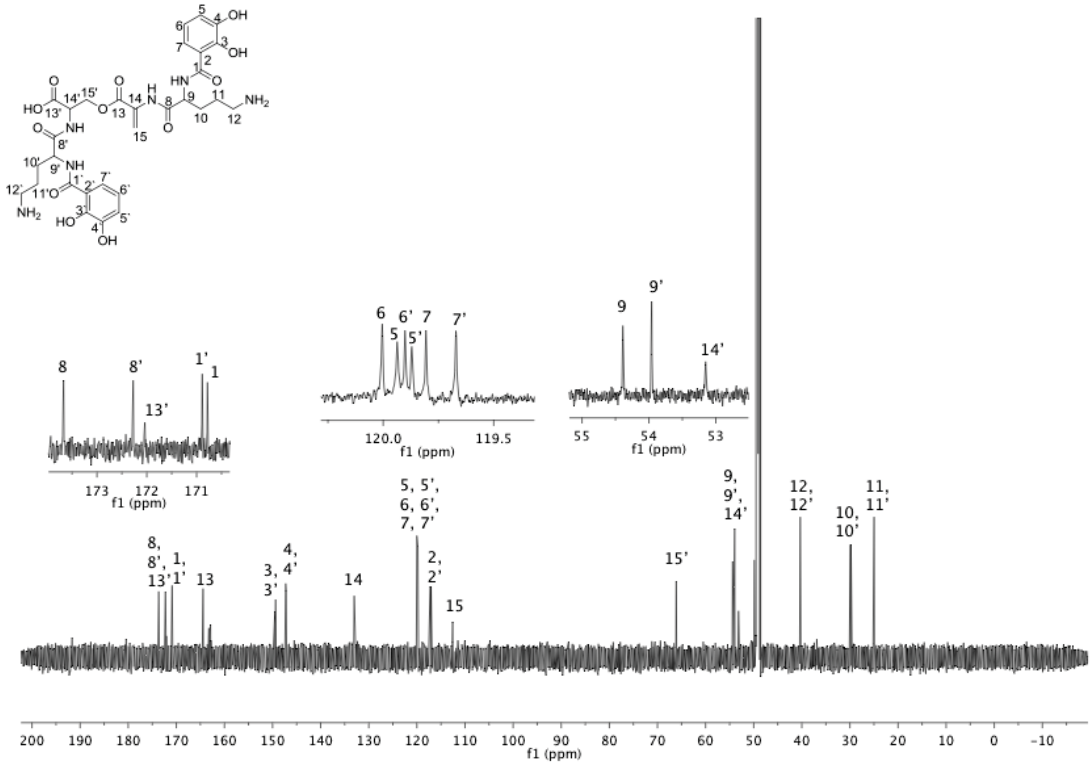

Figure A.25. Dehydrated diturnerbactin (**4**) ^{13}C NMR spectrum (800 MHz) in CD_3OD.

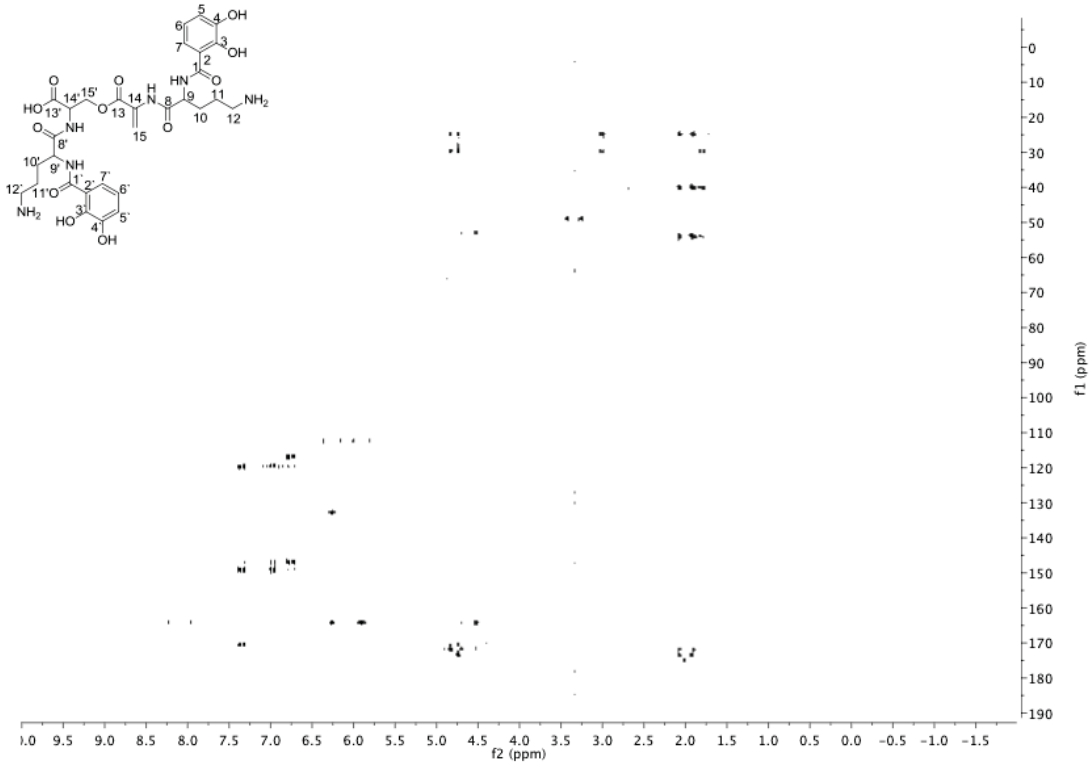

Figure A.26. Dehydrated diturnerbactin (**4**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD.

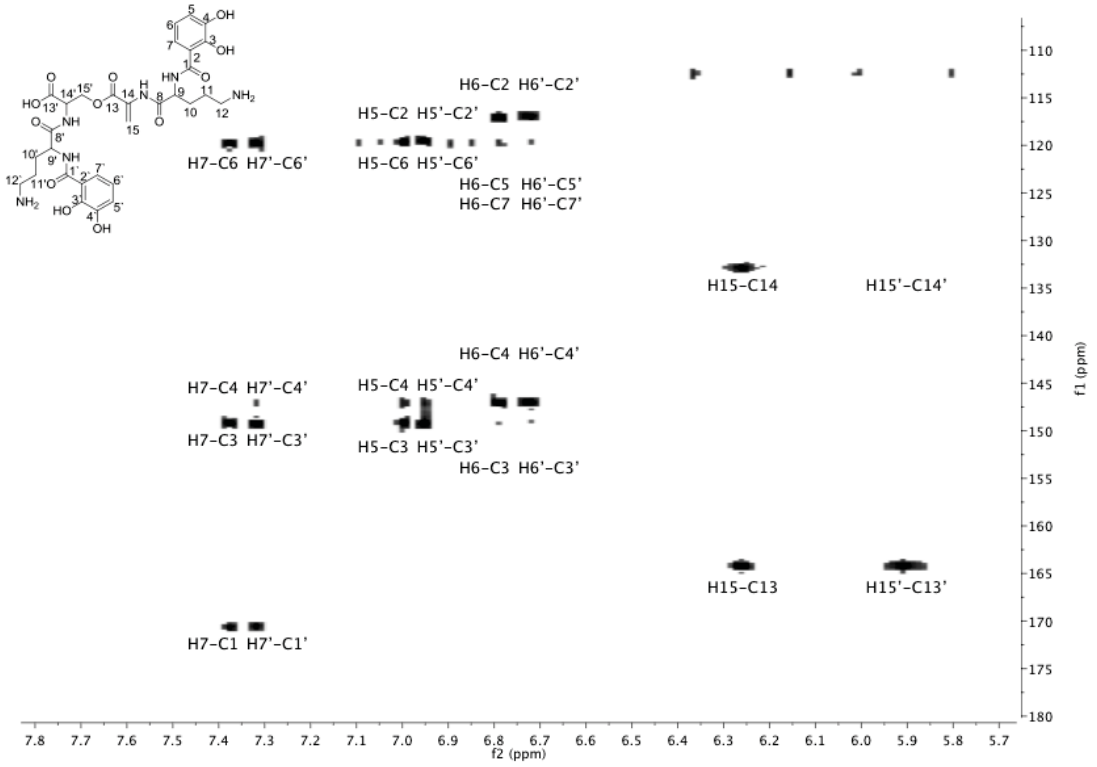

Figure A.27. Dehydrated diturnerbactin (**4**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

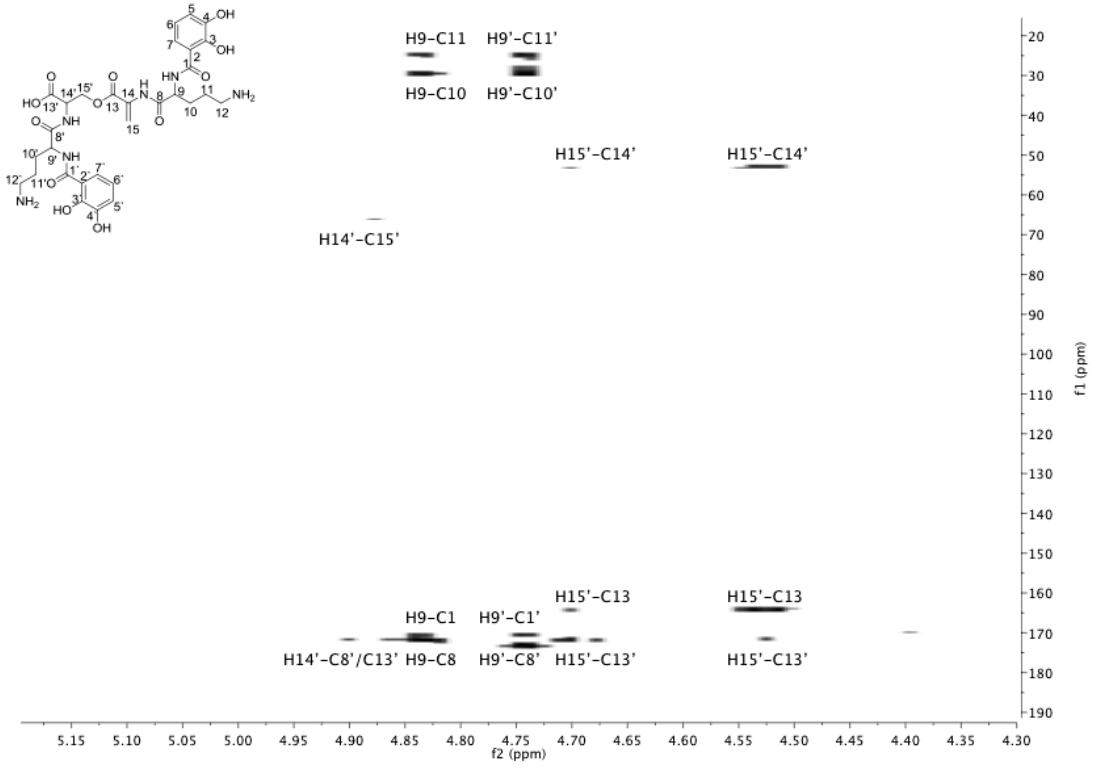

Figure A.28. Dehydrated diturnerbactin (**4**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

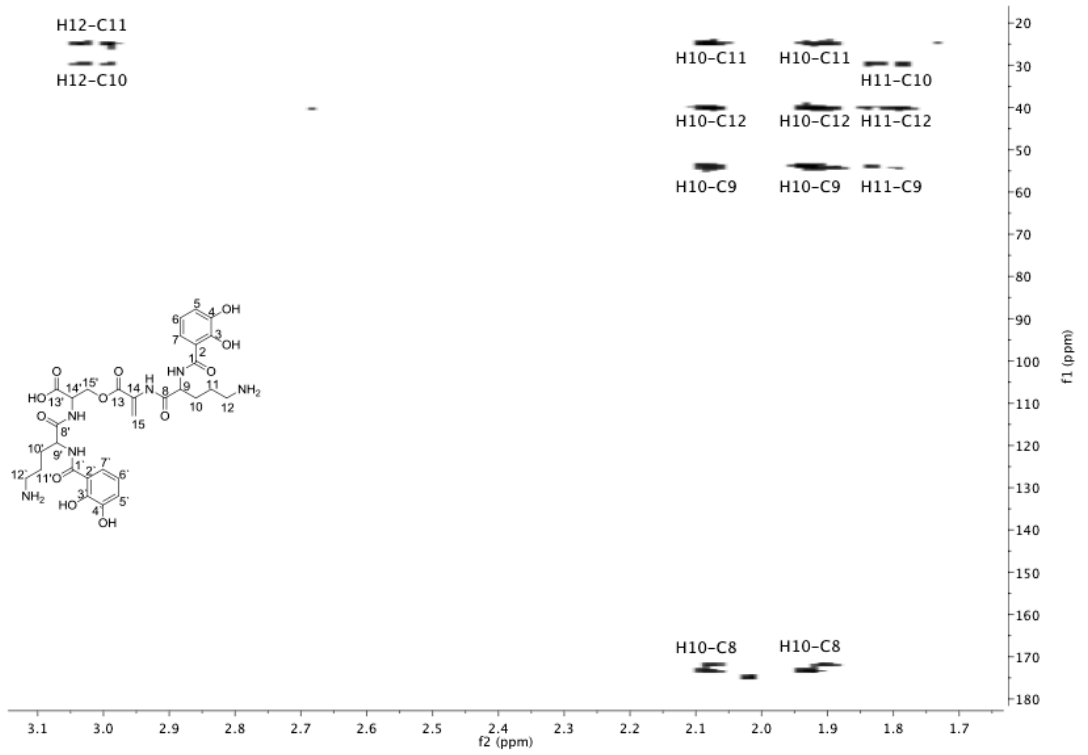

Figure A.29. Dehydrated diturnerbactin (**4**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

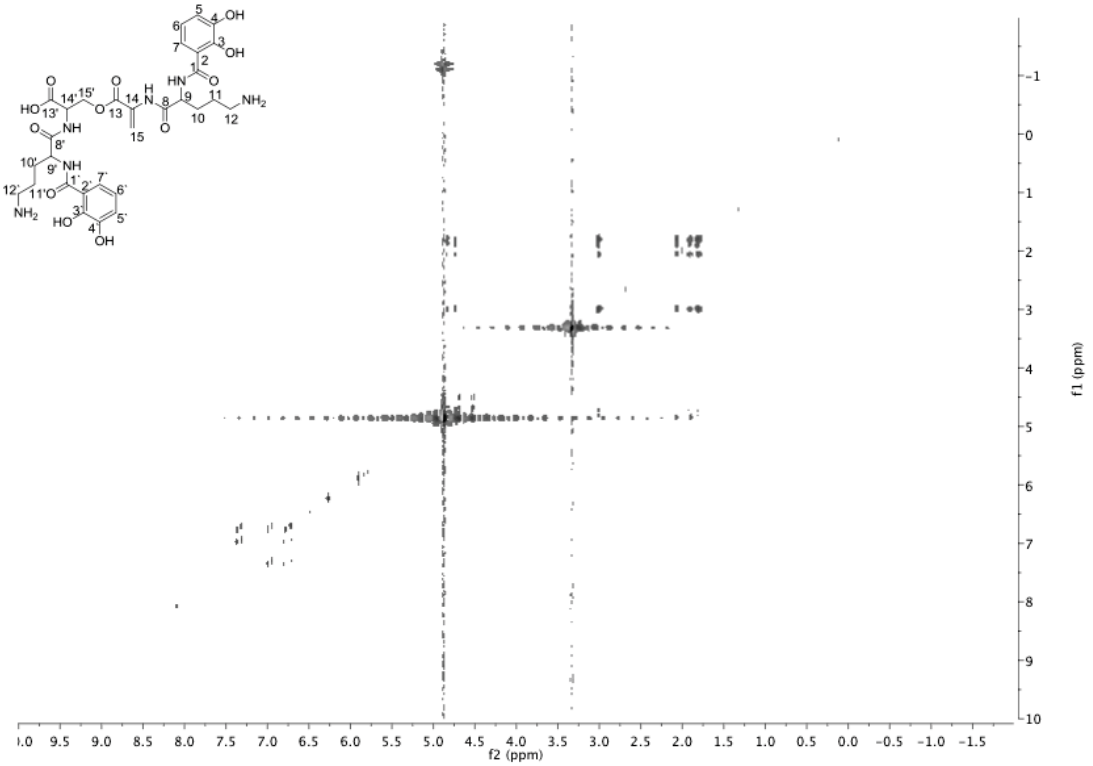

Figure A.30. Dehydrated diturnerbactin (**4**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD.

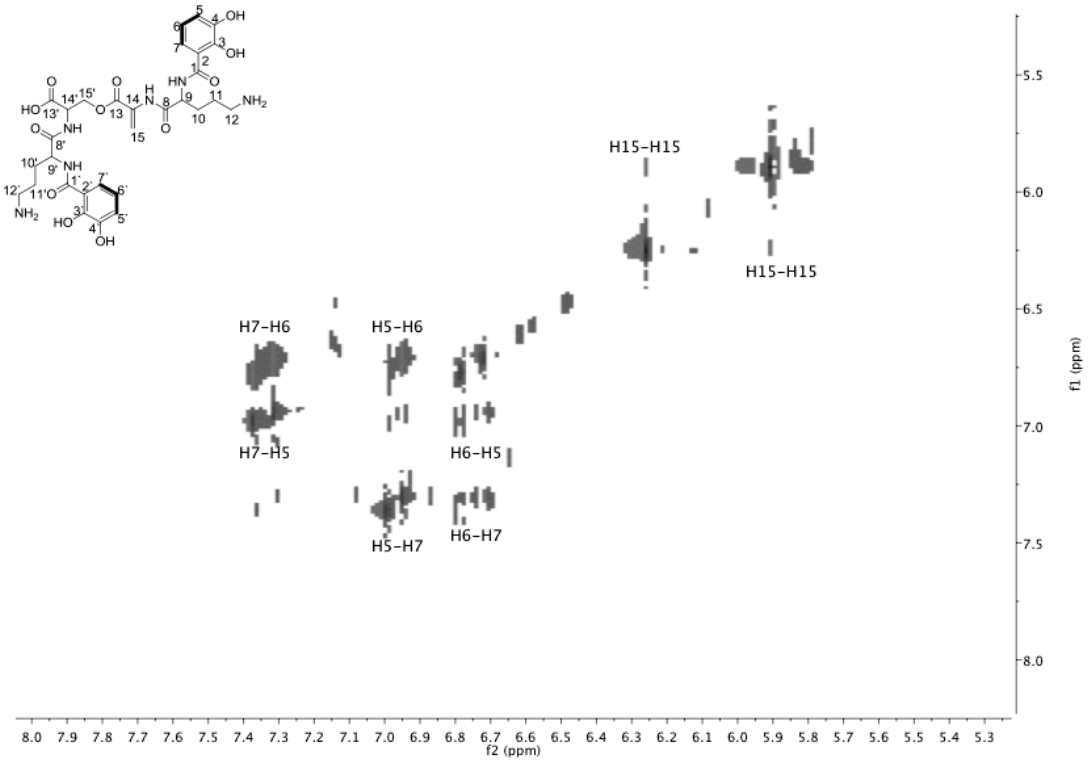

Figure A.31. Dehydrated diturnerbactin (**4**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

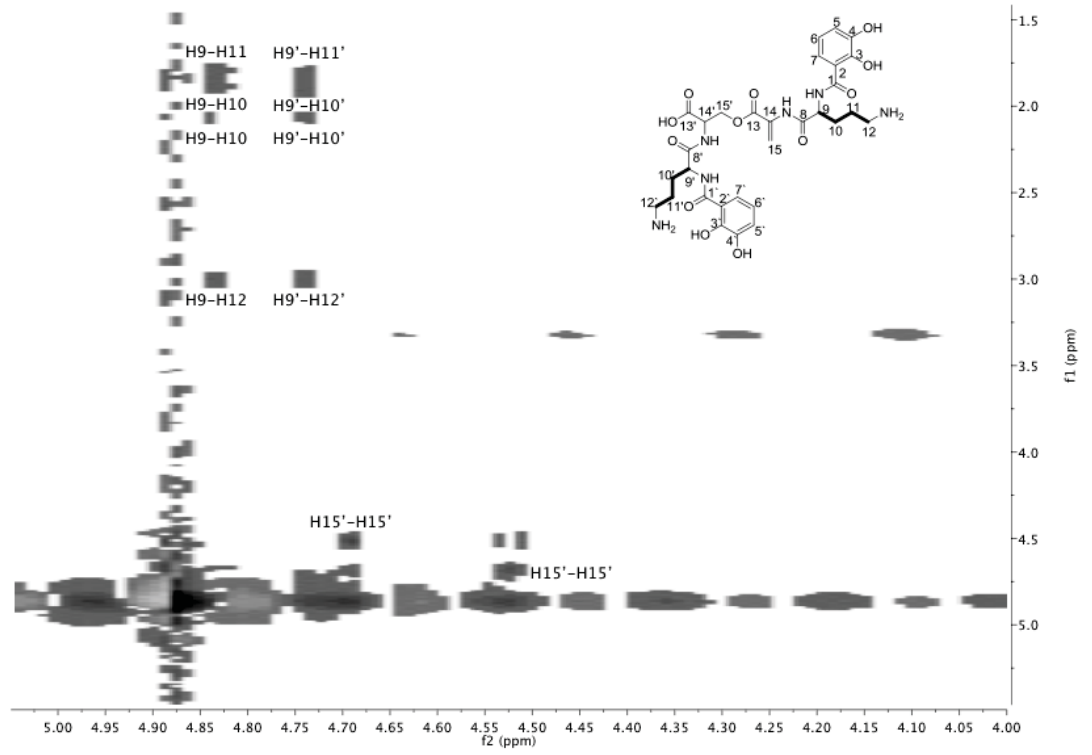

Figure A.32. Dehydrated diturnerbactin (**4**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

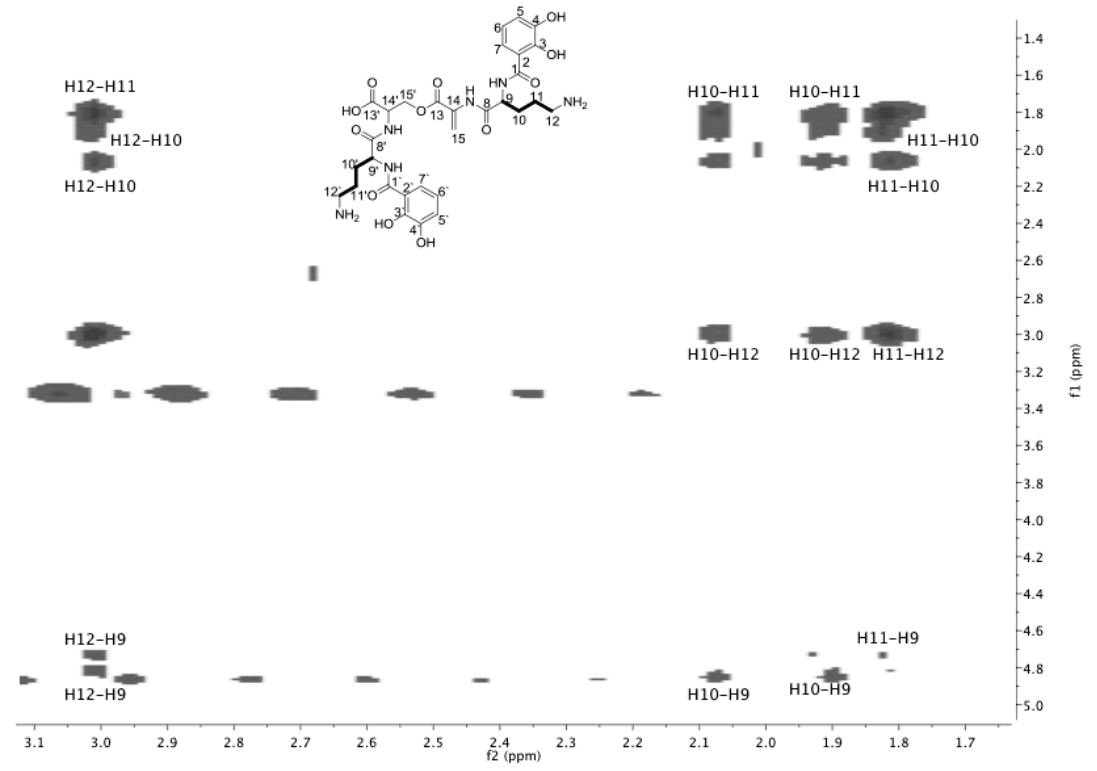

Figure A.33. Dehydrated diturnerbactin (**4**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

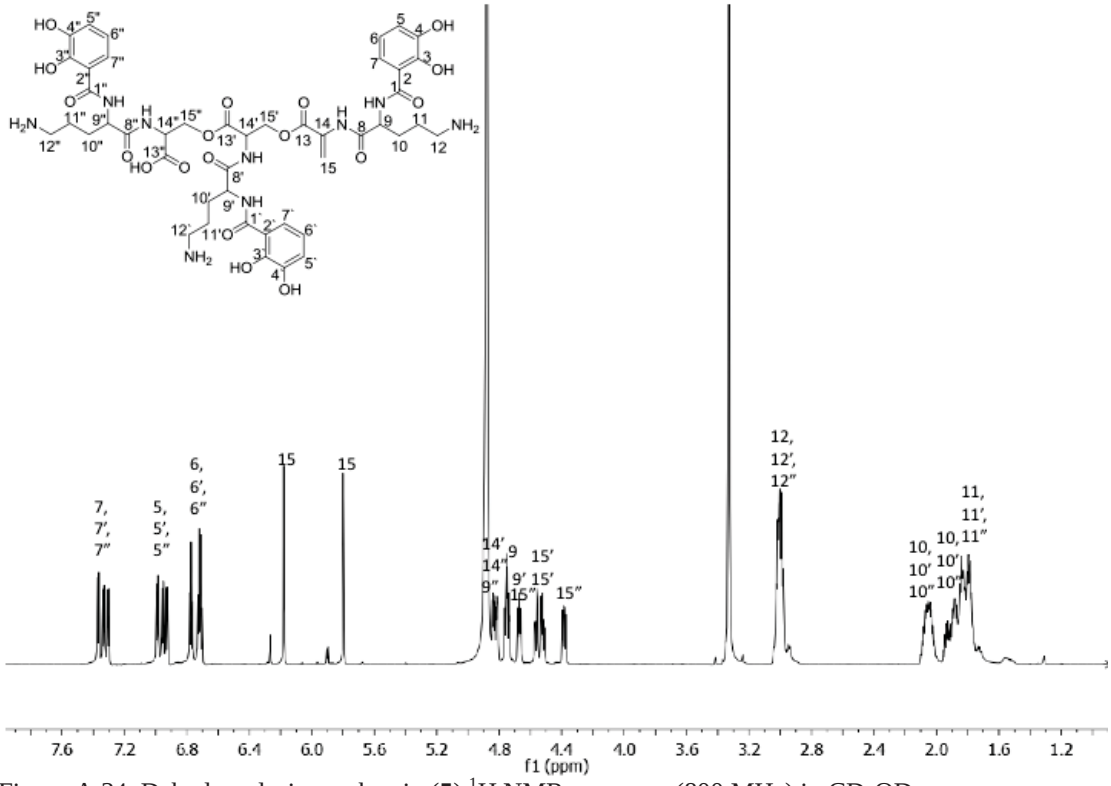

Figure A.34. Dehydrated triturnerbactin (**5**) ^{1}H NMR spectrum (800 MHz) in CD_3OD.

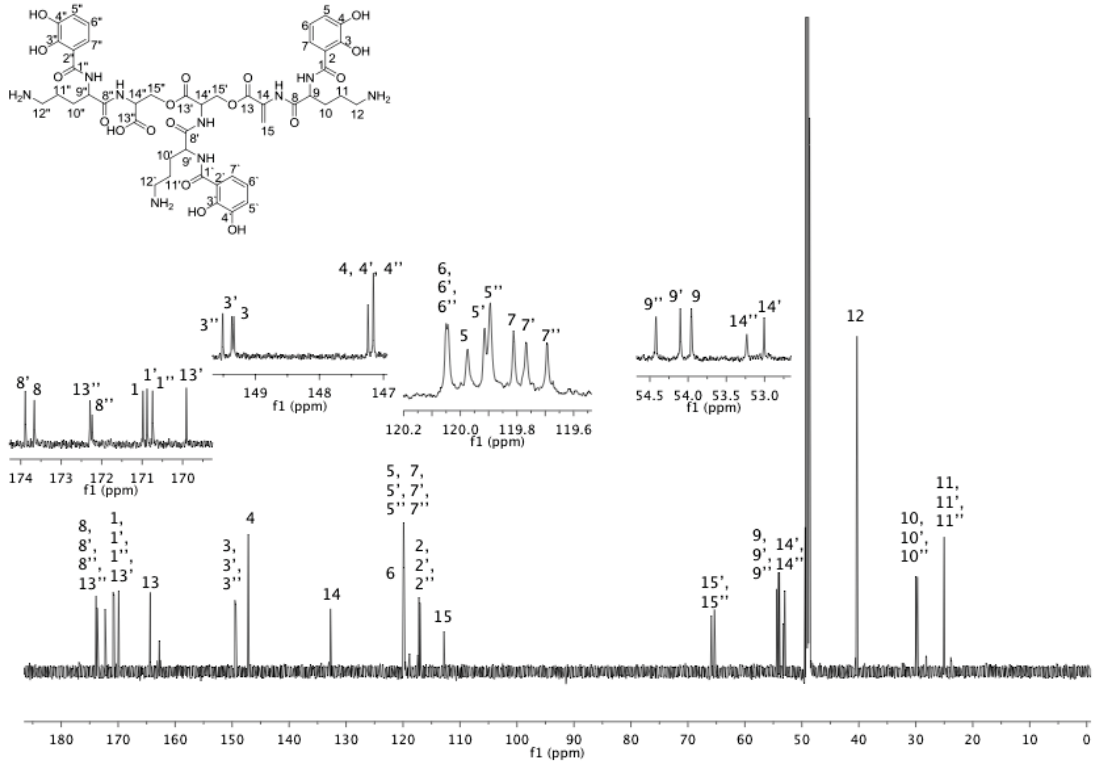

Figure A.35. Dehydrated triturnerbactin (**5**) ^{13}C NMR spectrum (800 MHz) in CD_3OD.

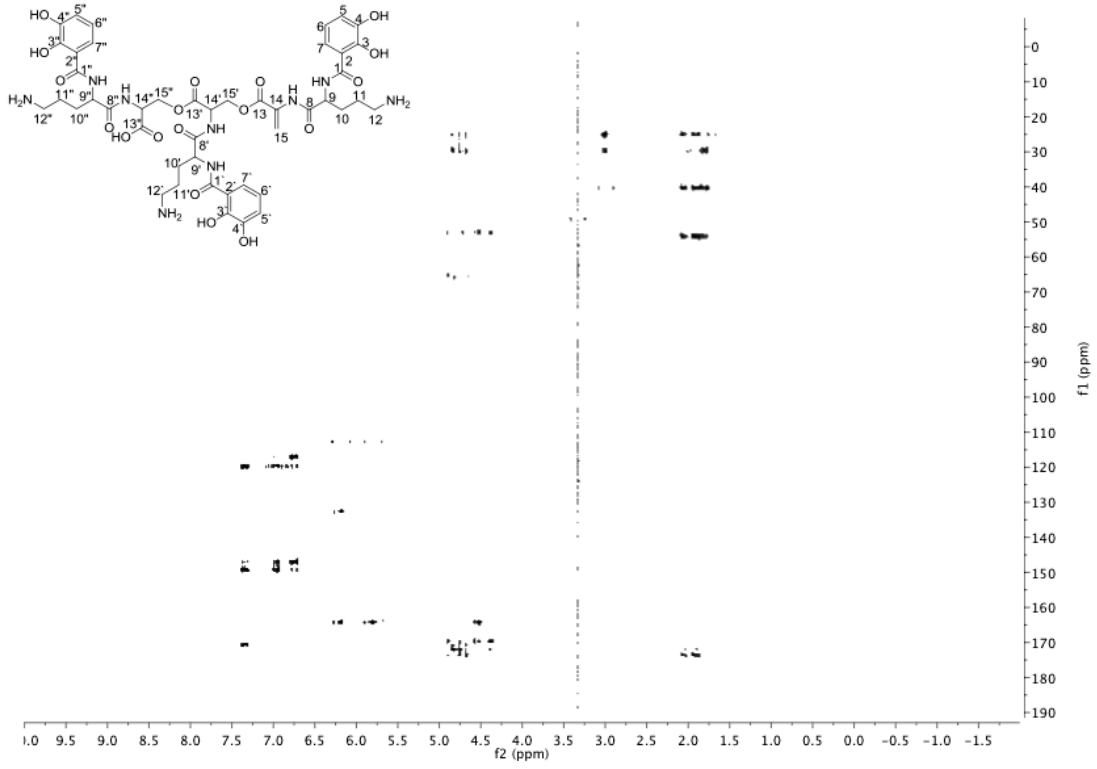

Figure A.36. Dehydrated triturnerbactin (**5**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD.

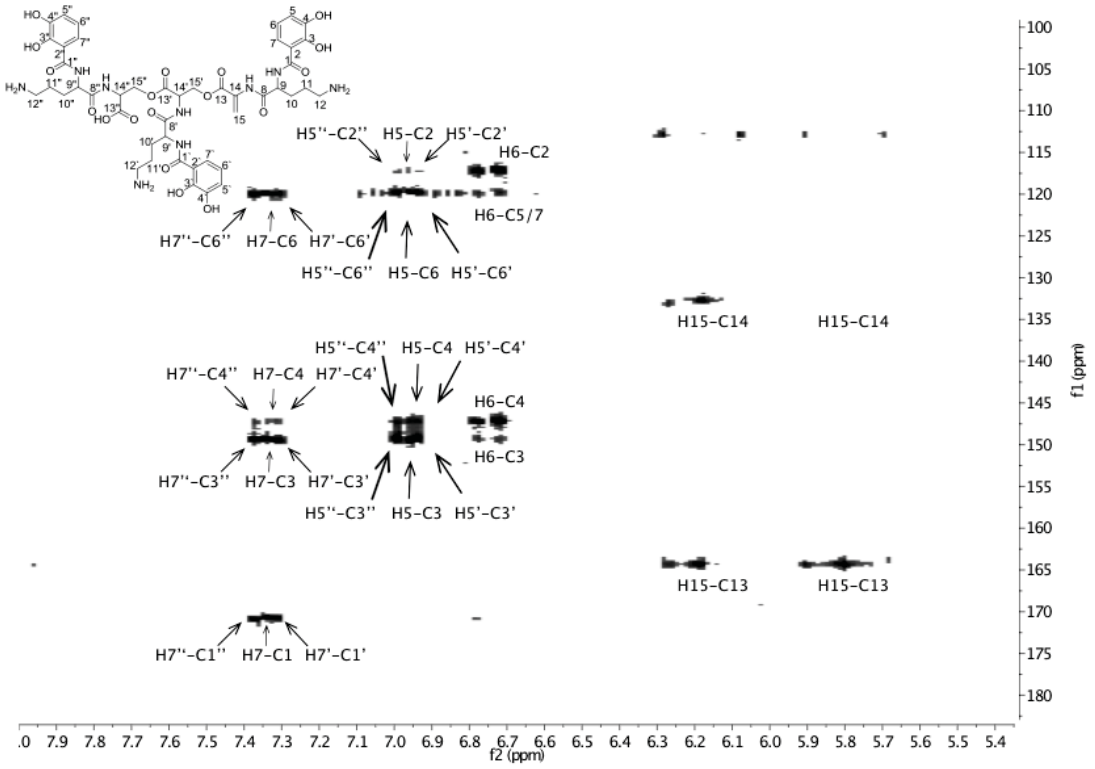

Figure A.37. Dehydrated triturnerbactin (**5**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

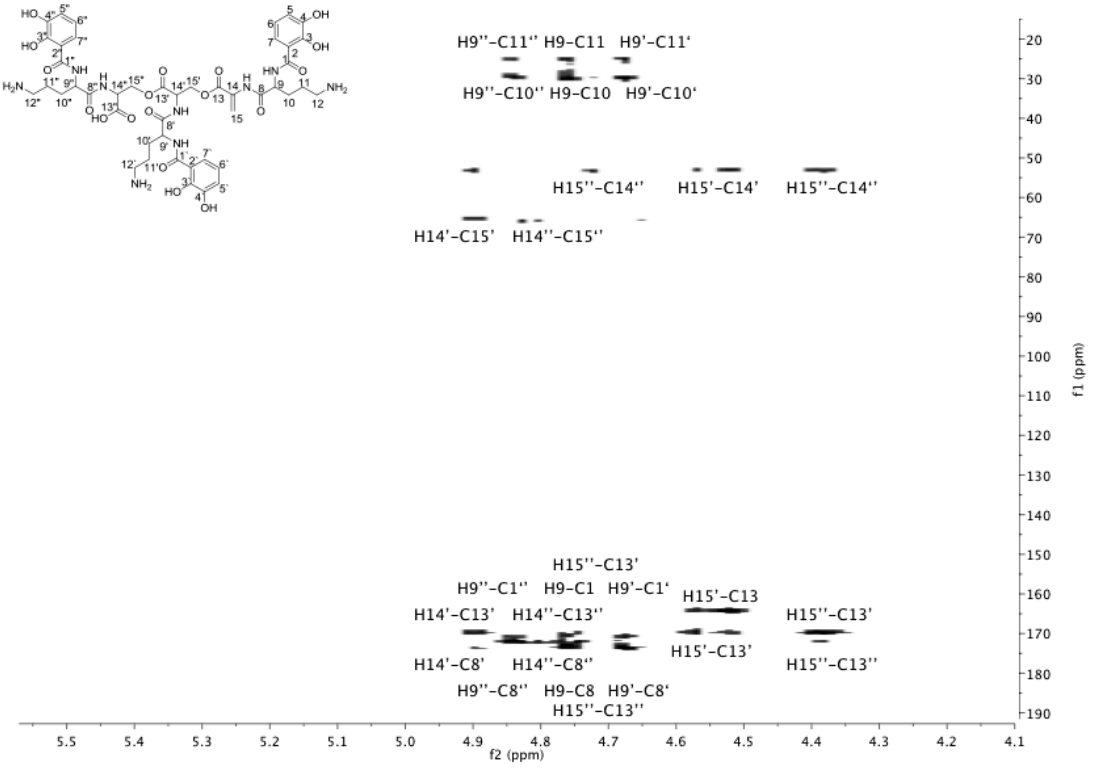

Figure A.38. Dehydrated triturnerbactin (**5**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

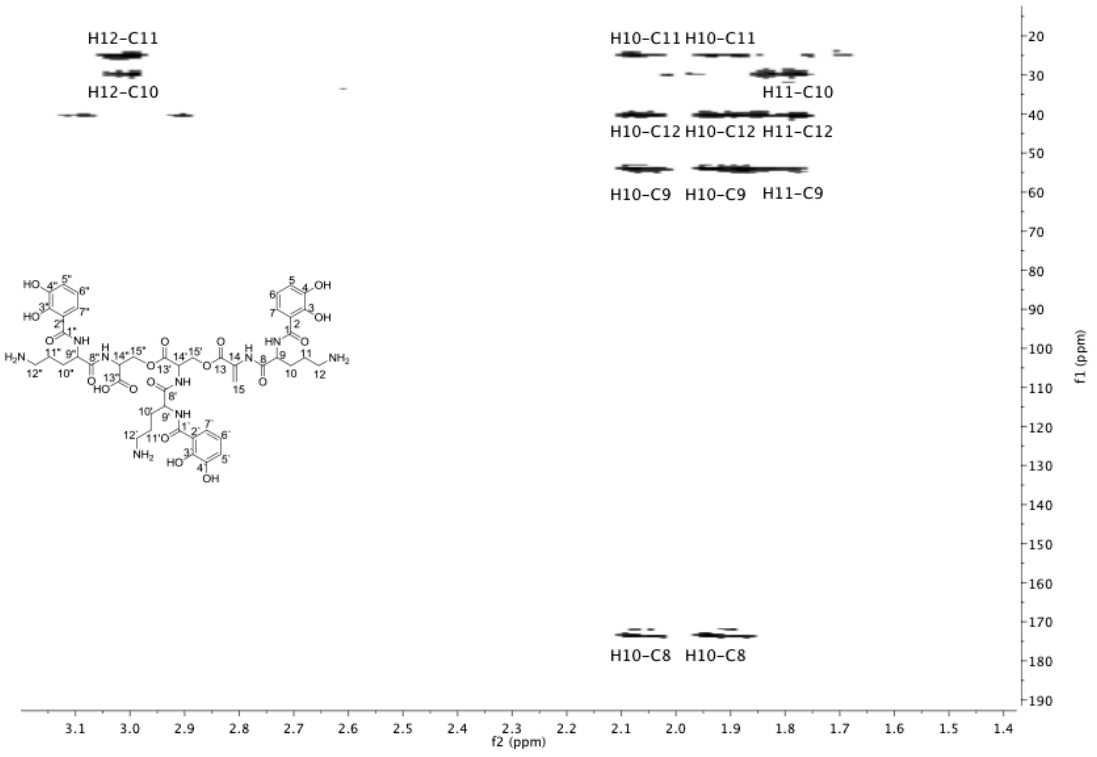

Figure A.39. Dehydrated triturnerbactin (**5**) ^{1}H-^{13}C HMBC spectrum (800 MHz) in CD_3OD, expanded region.

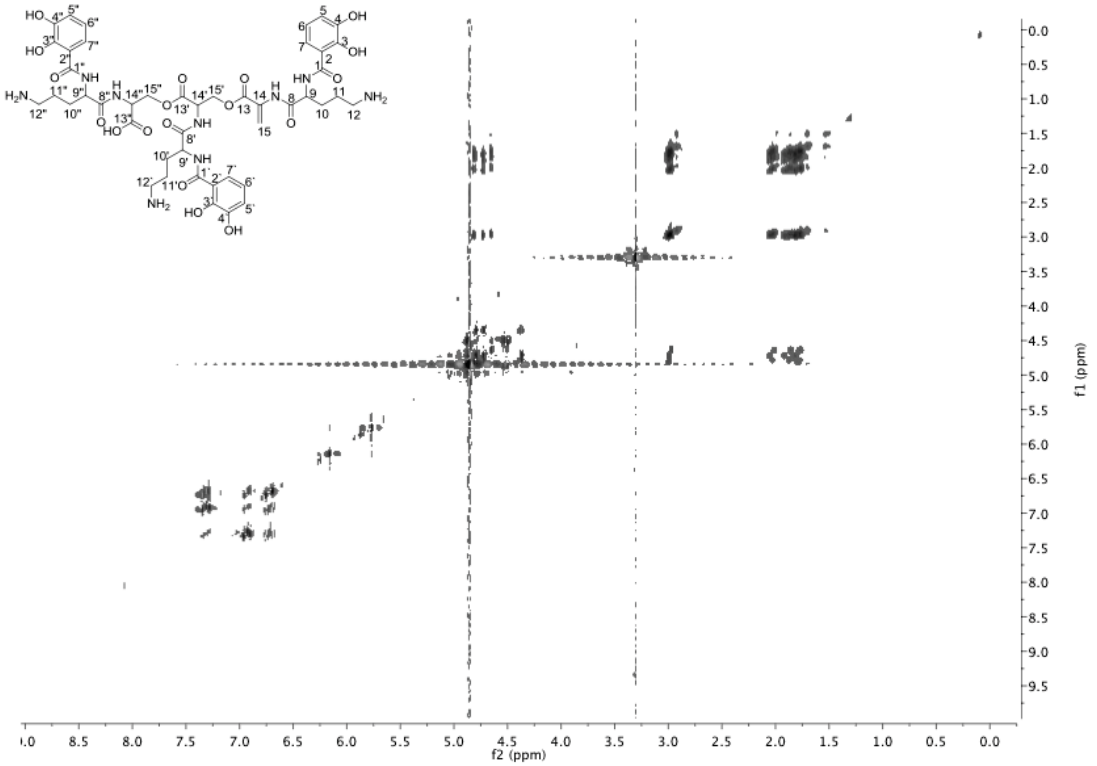

Figure A.40. Dehydrated triturnerbactin (**5**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD.

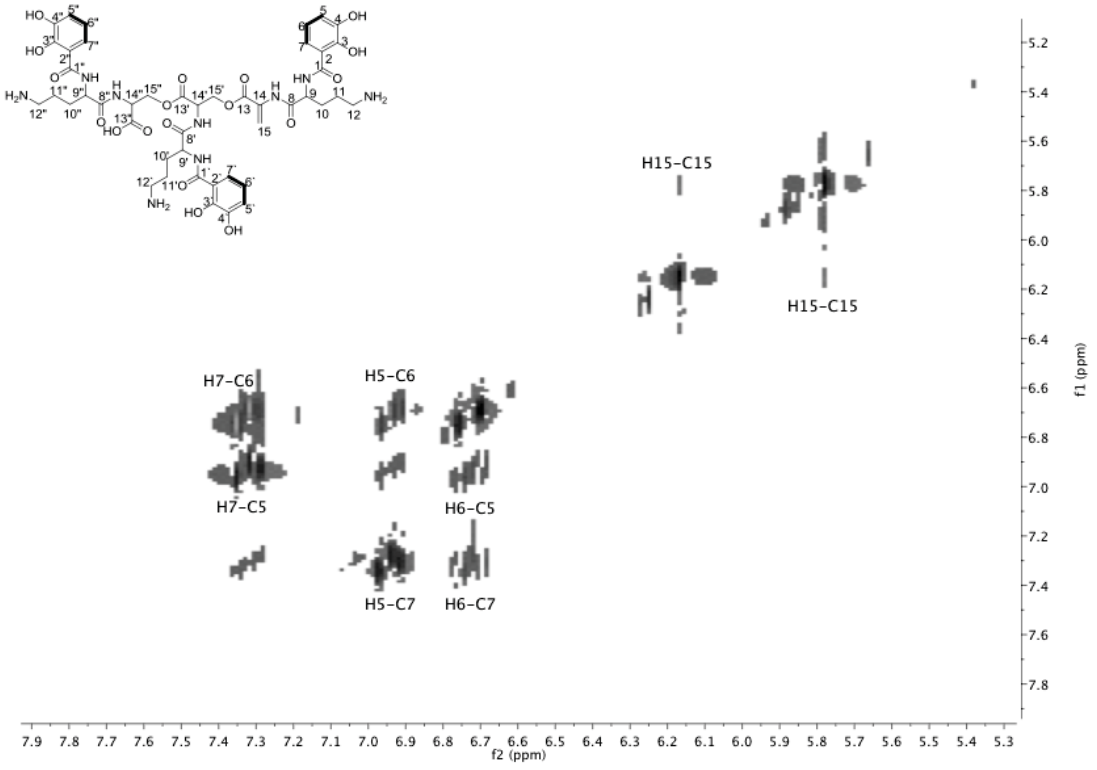

Figure A.41. Dehydrated triturnerbactin (**5**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

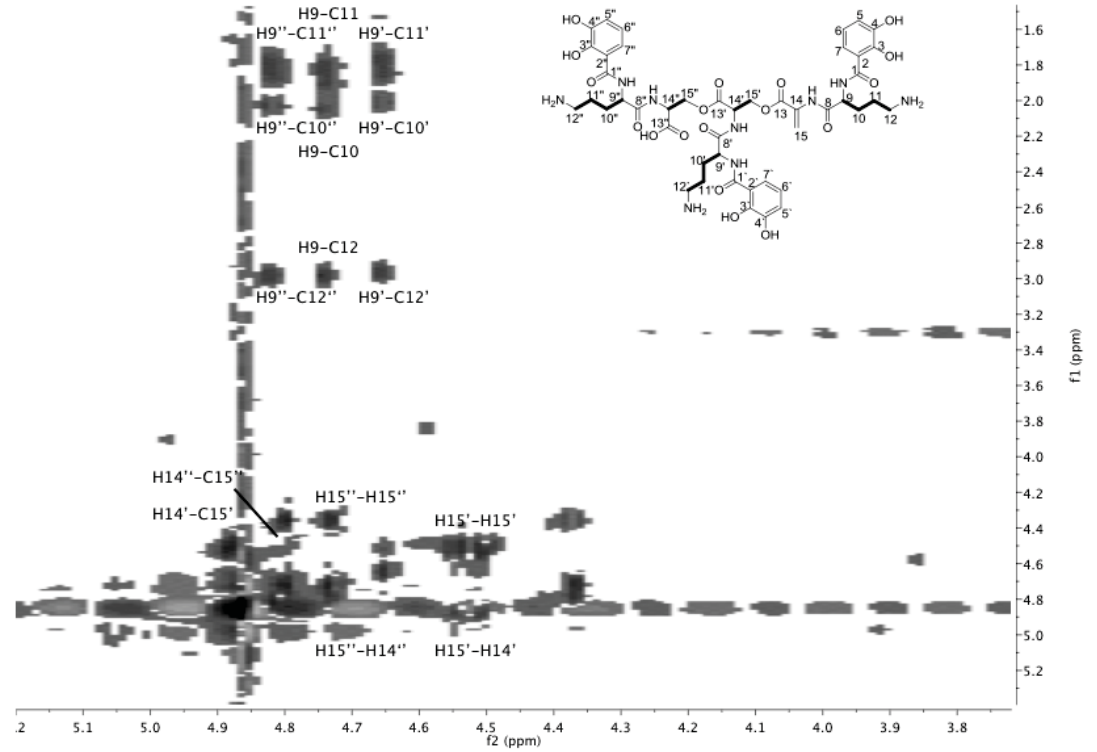

Figure A.42. Dehydrated triturnerbactin (**5**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

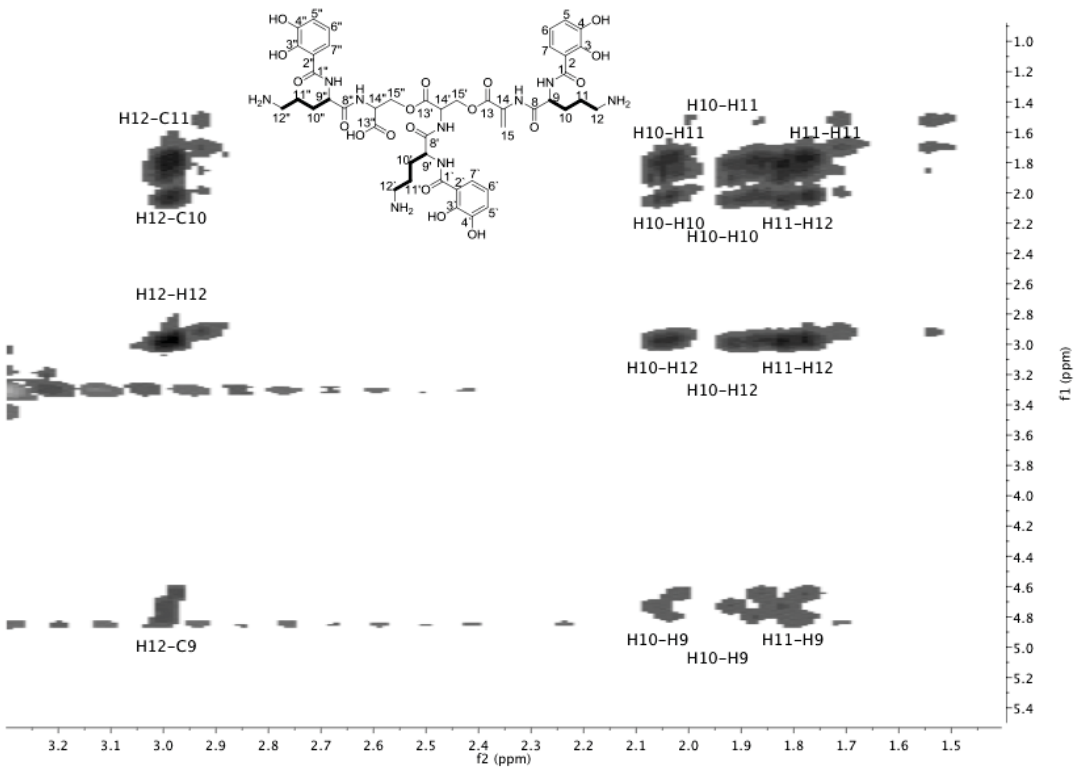

Figure A.43. Dehydrated triturnerbactin (**5**) ^{1}H-^{1}H TOCSY spectrum (800 MHz) in CD_3OD, expanded region.

Appendix B. The phylogeny of *Vibrio* sp. DS40M4 and an analysis of its siderophore biosynthetic machinery

B.1 Introduction

The open ocean isolate *Vibrio* sp. DS40M4 was found to produce a new triscatechol amide siderophore trivanchrobactin, the related biscatecholamide divanchrobactin, and the known siderophore vanchrobactin, as well as the known siderophore anguibactin. The known siderophores, vanchrobactin and anguibactin, had been previously described from strains of the fish pathogen *Vibrio anguillarum* (Actis et al., 1986; Soengas et al., 2006). However, both siderophores had not been previously isolated from the same strain, nor from a species other than *V. anguillarum*.

This work presents a 16S rRNA gene phylogenetic analysis of *Vibrio* sp. DS40M4. In addition, a bioinformatic analysis of the putative TE domain involved in vanchrobactin biosynthesis in DS40M4 is presented that supports the release of a linear peptide product. The phylogenetic analysis of DS40M4 appears in the following publication: Sandy, M., A. Han, J. Blunt, M. Munro, M. Haygood, A. Butler. 2010. Vanchrobactin and anguibactin siderophores produced by *Vibrio* sp. DS40M4. *J. Nat. Prod*. 73, 1038–1043.

B.2 Methods

B.2.1 Bacterial Strain

Vibrio sp. DS40M4 was isolated from an open ocean water sample which was collected over the continental slope off the West Coast of Africa between the Cape Verde and Canary Islands at 20° 41.1'N, 24° 13.7'W (Haygood et al., 1993).

B.2.2 Genomic DNA isolation

Vibrio sp. DS40M4 was grown in an artificial seawater medium (2 L) containing 10 g/L casamino acids, 19 mM NH_4Cl, 4.6 mM sodium glycerophosphate hydrate, 50 mM $MgSO_4$, 10 mM $CaCl_2$, 0.3 M NaCl, 10 mM KCl, 41 mM glycerol, 10 mM pH 7.4 HEPES buffer, 2 mM $NaHCO_3$, 8.2 μM biotin, 1.6 μM niacin, 0.33 μM thiamin, 1.46 μM 4-aminobenzoic acid, 0.21 μM panthothenic acid, 5 μM pyridoxine hydrochloride, 0.07 μM cyanocobalamin, 0.5 μM riboflavin, and 0.5 μM folic acid. One mL of overnight culture was harvested by centrifugation at 11000 rpm for 3 min. Genomic DNA from the cell pellet was isolated using the DNeasy Kit (Qiagen) following the bacterial DNA purification protocol.

B.2.3 Amplification of the PCP-TE domain of the vanchrobactin NRPS

Degenerate primer design: A multiple sequence alignment of PCP-TE didomains from NRPS gene clusters related to *V. anguillarum* RV22's vanchrobactin biosynthesis NRPS was constructed. The following PCP-TE didomains were used: *V. anguillarum* RV22 vabF (CAJ45639), *Escherichia coli* entF (ABV17231), *Salmonella enterica* entF (ZP_03366305), *Klebsiella pneumoniae* entF (ACI11302),

Erwinia chrysanthemi cbsF (AAL66867), *Bacillus subtilis* dhbF (CAB15186). PCP-TE didomains were aligned using the protein multiple sequence alignment program, MUSCLE (Edgar, 2004). This alignment was then imported into CODEHOP (Rose et al., 2003) and degenerate primers were designed using the codon usage table for *V. cholera* and *V. fischeri*. The following degenerate primers were designed and used in this work: PCP-TEF: 5'-GCRAATGGTAAAYTRGATAARAARGCNYTNCC-3', and PCP-TER: 5'-CCACCRAAAGACTAACCYAAYAR-3'.

PCR amplification: PCR with degenerate PCP-TE didomain primers was performed using genomic DNA as a template. In order to find a suitable annealing temperature, gradient PCR was carried out from a range of 45° C to 60° C. The highest annealing temperature at which bands were visualized on an agarose gel was used moving forward. Subsequent amplifications were then carried with the following profile: initial denaturation for 2 min at 94° C, 40 cycles of 1 min at 94° C, 1 min at 47° C, and 1 min at 72° C, and a final extension step for 7 min at 72° C. PCR products were analyzed by electrophoresis on a 1.2% agarose gel, and bands corresponding to the expected size were cut out and purified using the GENECLEAN II Kit (MP Biomedicals). Purified PCR products were cloned into the pCR4 vector using the TOPO TA Cloning Kit for Sequencing (Invitrogen). Plasmids were extracted using the QIAprep Spin Miniprep Kit (Qiagen). Inserts were sequenced with two-fold coverage using M13F and M13R primers on an ABI3700XL at the Molecular and Cellular Biology Core on OHSU's West Campus (Beaverton, OR). Sequences were assembled in Sequencher 4.8 (GeneCodes Corp.). Corresponding amino acid sequences were searched against the National Center for Biotechnology Information

(NCBI) nonredundant protein database by using the Basic Local Alignment Search Tool (BLAST).

B.2.4 Phylogenetic Analysis

General bacterial 16S rRNA gene primers 27F and 1492R (Lane, 1991) were used to amplify the nearly full-length 16S rRNA gene from *Vibrio* sp. DS40M4 genomic DNA. PCR products were analyzed by electrophoresis on an agarose gel (1.2%) to confirm size and specificity. PCR products were cleaned using the QIAquick PCR purification kit (Qiagen) and then directly sequenced. Both strands of the PCR product were fully sequenced using internal primers (Lane, 1991). Sequencing was carried out on an ABI3700XL at the Molecular and Cellular Biology Core on OHSU's West Campus (Beaverton, OR). Sequences were compiled in Bioedit (Hall, 1999) resulting in a sequence of 1437 bp in length. The GenBank accession number is HM152762. The 16S rRNA gene sequence was compared to those in the NCBI nucleotide collection and RDP databases. The 16S rRNA sequence was aligned with the SINA Webaligner (SILVA) (Pruesse et al., 2007). Aligned type strain reference sequences were downloaded from SILVA's rRNA database. The multiple sequence alignment was compiled and edited in Bioedit. Hypervariable regions were excluded from the analysis. Phylogenetic reconstruction was performed on unambiguously aligned nucleotide positions with RAxML v. 7.0.4 (Stamatakis 2006) using the General Time Reversible model of nucleotide substitution under the Γ model of rate heterogeneity (GTRGAMMA) with 100 bootstrap replicates. The selected tree topology had the highest likelihood score out of

100 heuristic tree searches, each search beginning with a distinct randomized maximum parsimony starting tree.

B.2.5 Evaluation of alternative phylogenetic hypotheses

Alternative phylogenetic hypotheses for *Vibrio* sp. DS40M4 were evaluated by the approximately unbiased (AU) test (Shimodaira, 2002). Constraint trees were constructed placing *Vibrio* sp. DS40M4 in various clades appearing in the original unconstrained maximum likelihood tree. RAxML was used with the original data set to infer maximum likelihood phylogenies for each constraint under the GTRGAMMA model using the same parameters as the original search. Site-wise log likelihoods were estimated by RAxML under the GTRGAMMA model for each constraint tree topology. The AU test was implemented by CONSEL v. 0.1j (Shimodaira and Hasegawa, 2001).

B.3 Results

B.3.1 Phylogenetic Analysis of the Bacterial Small Subunit (16S) rRNA Gene

A BLAST search of GenBank using the 16S rRNA gene from *Vibrio* sp. DS40M4 revealed over 99% similarity to multiple previously described *Vibrio* strains, including *V. campbellii, V. rotiferianus*, and *V. harveyi*; the relatively conserved SSU rRNA gene has limited power for resolving this closely related group. A phylogenetic tree created by maximum likelihood reveals that *Vibrio* sp. DS40M4 forms a cluster with these strains (Figure B.1). *Vibrio* sp. DS40M4 did not, however,

cluster with *Vibrio* (*Listonella) anguillarum*, the only other bacterial species known to produce vanchrobactin and anguibactin.

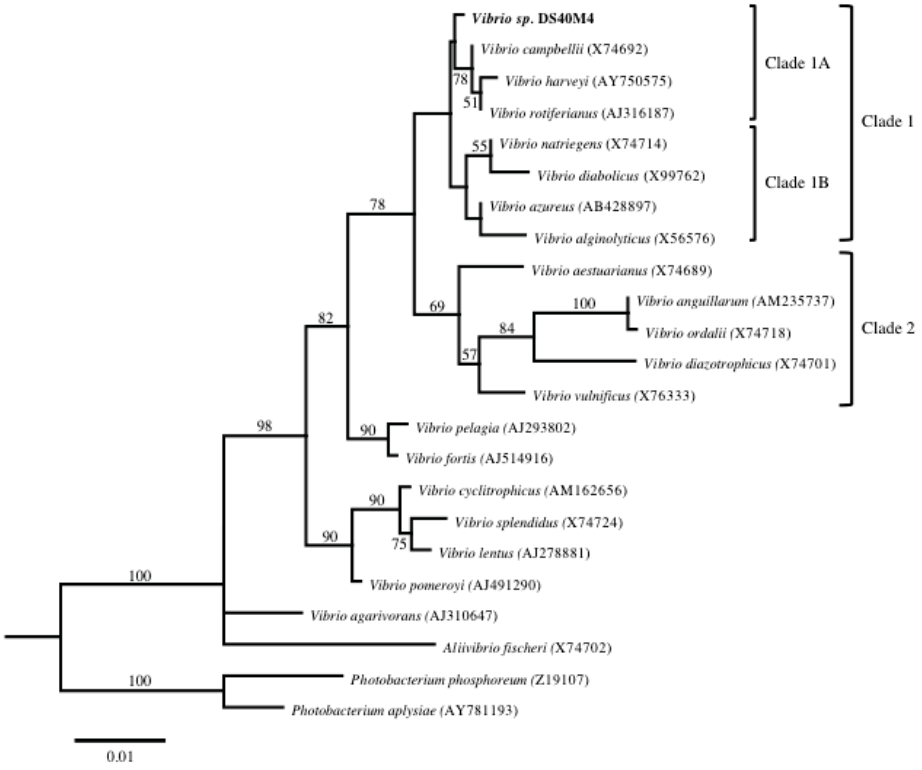

Figure B.1. Maximum-likelihood phylogenetic tree showing the placement of *Vibrio* sp. DS40M4 relative to previously described *Vibrio* strains. Bootstrap values over 50% are shown. Scale bar represents 0.01 nucleotide substitution per site.

B.3.2 Evaluation of Alternative Phylogenetic Hypotheses

Due to the low bootstrap values supporting the placement of *Vibrio* sp. DS40M4 within the *Vibrio* subgroup shown in Figure B.1, alternative hypotheses regarding its phylogenetic placement were tested using the AU test. Constraint trees placing *Vibrio* sp. DS40M4 in neighboring clades were constructed, and the results of that analysis are shown in Table B.1. In the original, best scoring likelihood tree, *Vibrio* sp. DS40M4 was placed in Clade 1A (Figure B.1). A topology

placing *Vibrio* sp. DS40M4 in Clade 1B, a sister group to Clade 1A, did not differ significantly from the most likely tree. However, alternative hypotheses placing *Vibrio* sp. DS40M4 in Clade 2, a clade with *V. anguillarum*, were rejected. These results show support for the placement of *Vibrio* sp. DS40M4 within Clade 1 and its exclusion from Clade 2.

Table B.1. Alternative Phylogenetic Hypotheses with Corresponding log-Likelihoods and *P* Values Inferred from the AU Test. Statistically significantly worse trees (rejection of the hypothesis) are those with a *P* value below 0.05 and are shown in bold.

Tree	- ln Likelihood	Difference from best tree	*P* value
Best Likelihood tree (unconstrained)	3103.685284	Best	
Constraint:			
Monophyly of DS40M4 with Clade 1B	3107.988337	4.303053	0.269
Monophyly DS40M4 with Clade 2	3151.702524	48.01724	**0.005**
Monophyly of DS40M4, *V. ordalii*, *V. anguillarum*	3200.726377	97.041093	**<0.001**
Monophyly of DS40M4 and *V. anguillarum*	3214.060006	110.374722	**<0.001**

B.3.3 Screening for Ent F Thioesterase Homolog in *Vibrio sp.* DS40M4

Degenerate PCR primers were designed to amplify the vanchrobactin NRPS PCP-TE didomain from *Vibrio* sp. DS40M4 based on conserved amino acid sequence motifs, targeting the N-terminus of the PCP domain and the conserved GxSxG signature motif of the TE domain. PCR with these primers resulted in a product of expected size (~600 bp), as well as some non-specific products. The product of expected size was cut out of agarose gels, purified, cloned and sequenced.

The deduced amino acid sequence was searched against the NCBI's protein database using their BLASTP algorithm. The sequence obtained from *Vibrio* sp.

DS40M4 had high similarity to the NRPS responsible for vanchrobactin biosynthesis from *V. anguillarum* serovar O2 (74% identity over 194 aligned amino acids). Through α/β-hydrolase family sequence alignments and mutational analysis, a conserved Pro residue near the oxyanion hole of the NRPS TE active site has been shown to drive the mechanism of TE catalysis to cyclization over hydrolysis. Alignment of the putative TE domain of *Vibrio* sp. DS40M4's vanchrobactin biosynthesis gene cluster to a multiple protein sequence alignment of other NRPS TE domains shows that the putative TE domain of *Vibrio* sp. DS40M4 lacks this particular Pro, suggesting that the TE domain catalyzes a hydrolytic release of its product (Figure B.2).

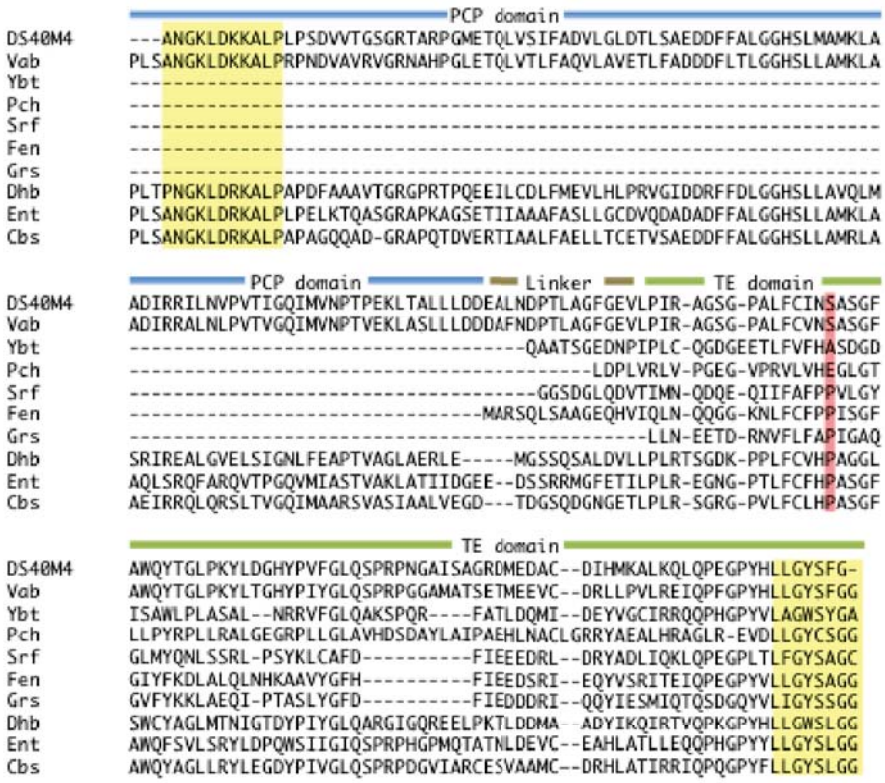

Figure B.2. Sequence alignment of the putative TE domain from *Vibrio* sp. DS40M4's siderophore biosynthesis NRPS. Shaded regions in yellow represent sequence motifs targeted with degenerate primers. Shaded region in red represents the site of the conserved Pro shown to influence cyclization. The sequences of PCP domains used in degenerate primer design are shown. Abbreviations are as follows, including corresponding accession numbers in parentheses: Vab, vanchrobactin, (CAJ45639); Ybt, yersiniabactin, (AAS61883); Pch, pyochelin, (AAG07613); Srf, surfactin, (2VSQ_A); Fen, fengycin, (2CB9_A); Grs, gramicidin S, (BAA06146); Dhb, bacillibactin, (CAB15186); Ent, enterobactin, (ABV17231); Cbs, chrysobactin, (AAL66867).

B.4 Discussion

Vanchrobactin and anguibactin are known siderophores produced by various strains of *Vibrio anguillarum*, a fish pathogen causing vibriosis; however they have never before been isolated from the same strain. Vanchrobactin has been isolated from *V. anguillarum* serotype O2 strain RV22 (Soengas et al., 2006) whereas anguibactin has been isolated from *V. anguillarum* serotype O1 strain 775 (pJM1) (Actis et al., 1986). On the basis of phylogenetic analysis of the SSU rRNA gene, DS40M4 is a *Vibrio* sp. strain that falls within the *Vibrio campbellii* group; it is

clearly not a strain of *V. anguillarum* on the basis of the fact that its sequence does not group with that of the *V. anguillarum* type strain (Figure B.1). This work expands the distribution of both vanchrobactin and anguibactin to *Vibrio* species other than *V. anguillarum*.

The TE domain homologue recovered from DS40M4 is likely to originate in the vanchrobactin biosynthesis pathway, based on its high similarity to the vanchrobactin TE of *V. anguillarum* serotype O2 strain RV22, and dissimilarity to that of the anguibactin TE (Less than 20% identity over 94 aligned amino acids). Sequence analysis of the putative TE domain of DS40M4's vanchrobactin biosynthesis gene cluster suggests that the siderophore should be a linear peptide product, as opposed to a cyclic product. This bioinformatic support is based on the absence of a conserved Pro that has been shown to drive catalysis by the TE domain to cyclization over hydrolysis (Tseng et al., 2002). This Pro is thought to impose a conformational constraint on the TE domain such that the peptidyl ester chain is protected from hydrolytic attack by a water molecule, allowing for intramolecular nucleophilic attack and cyclization. When this Pro was mutated to a Gly in SrfTE, the purported flexibility allowed by the Gly residue resulted in a significant shift towards hydrolytic release of the linear peptide product compared to the wild-type. Support for this hypothesis came when the P26G mutant SrfTE yielded only the cyclized product when the *in vitro* reaction took place in an organic solvent, N,N-dimethylformamide, establishing that the P26G mutant retains wild-type function when water is excluded from the reaction (Wagner et al., 2006).

Oligomerization by NRPSs is not a novel concept as established, for example, in the biosynthesis of enterobactin (Shaw-Reid et al., 1999; Frueh et al., 2008), bacillibactin (May et al., 2001), gramicidin S (Hori et al., 1989). The synthesis of oligomeric peptide products occurs through the action of iterative NRPS's, whose modules and/or domains are utilized more than once in the synthesis of a single product (Mootz et al., 2002). In contrast, some NRPSs function linearly in which modules are utilized only once during synthesis. Because we do not currently have the entire biosynthetic pathway for the trivanchrobactin siderophore from DS40M4, we cannot definitively state that the TE domain acts in an iterative manner. However, the appearance of a trimer suggests an iterative mode of action. The current finding of a linear trimer of a previously described monomer raises the question as to what determines whether a TE acts iteratively or linearly. Multiple sequence alignments of TE domains have not revealed any signature sequence motifs that will allow the prediction of either mode of action (Mootz et al., 2002), *in vitro* experiments (Hoyer et al., 2007) and structural analyses of TE domains (Samel et al., 2006) have explored this issue, but the factors that lead to an iterative-acting TE domain remain unknown.

Appendix C. The phylogenetic diversity of microbes associated with colonies of *Trichodesmium*

C.1 Introduction

Trichodesmium is a globally important nitrogen fixing cyanobacterium, contributing significant amounts of new nitrogen to the tropical and subtropical oceans. *Trichodesmium* often exists in colonies, which can provide an inviting habitat for marine microorganisms by providing a solid substrate and a source of nitrogen and carbon. The associated microbial community of *Trichodesmium* has been previously characterized, but a comparative community diversity survey of *Trichodesmium* colonies from different geographic locations has not been carried out.

The overall goal of this work was to examine *Trichodesmium* colonies collected from different geographic locations and identify the diversity of associated microorganisms, particularly searching for potential mutualistic relationships between *Trichodesmium* and its associated microorganisms. The results show the presence of *Alpha*, *Beta* and *Gammaproteobacteria* as well as members of the *Bacteroides* lineage associated with *Trichodesmium* colonies. A single highly related clade of *Alphaproteobacteria* was detected in *Trichodesmium* colonies from the Sargasso Sea, Pacific Ocean and Caribbean Sea that may represent a novel Order and have a specific, perhaps mutualistic relationship with *Trichodesmium* colonies.

The current chapter presents a subset of work contributing to the overall goals of the project mentioned above, led by Elizabeth Mann, currently an assistant professor at the Skidaway Institute of Oceanography. This work includes the phylogenetic analysis of *Betaproteobacteria* and *Gammaproteobacteria* associated

with *Trichodesmium* colonies. Additionally, the abundance of the unique *Trichodesmium* associated *Alphaproteobacteria* (TaA) clade is examined by a quantitative PCR assay.

C.2 Methods

C.2.1 Sample and data collection

The collection of *Trichodesmium* samples, extraction of DNA, amplification of 16S rRNA gene sequences, and DNA sequencing was performed by Elizabeth Mann. Briefly, *Trichodesmium* colonies were collected from the Sargasso Sea, the Caribbean Sea, and the Pacific Ocean using a 300 μm mesh plankton net. Samples were not separated by species and some samples contained multiple species of *Trichodesmium*. Colonies were rinsed five times with filter sterilized seawater and then were stored in a high sucrose lysis buffer (0.75 M sucrose, 50 mM Tris, 40 mM EDTA at pH 8.3). A differential lysis and DNA extraction protocol was developed in order to separate *Trichodesmium* DNA from that of associated microbes. *Trichodesmium* was mechanically lysed by pipetting and vortexing, then centrifuged for 1.5 hours at 14000 rpm to separate the *Trichodesmium* DNA from the pellet of intact associated bacterial cells. This pellet was rinsed, treated with DNase to remove any residual *Trichodesmium* DNA, and then the pellet DNA was extracted using a standard phenol/chloroform procedure with ethanol precipitation. 16S rRNA gene sequences were amplified with two sets of primer sets: 27F and 1492R, 530F and 1392R. Amplicons were cloned using the TOPO TA cloning kit (Invitrogen) and then sequenced. Sequences were checked for chimeras using the RDP Chimera

Check program, Bellerophon, and Pintail. Operational taxonomic units (OTUs) and diversity assessments were calculated using DOTUR.

C.2.2 Phylogenetic analysis

Trichodesmium-associated 16S rRNA gene sequences were aligned with the SILVA's SINA Webaligner (Pruesse et al., 2007). Aligned reference sequences were downloaded from SILVA's rRNA database project (http://www.arb-silva.de). Sequence alignments were compiled and edited in Bioedit (Hall, 1999). Hypervariable regions were removed from the analysis. Phylogenetic reconstructions were performed using RAxML v. 7.0.4 (Stamatakis, 2006) using the General Time Reversible model of nucleotide substitution under the Γ model of rate heterogeneity (GTRGAMMA) with 100 bootstrap replicates. The selected tree topology had the highest likelihood score out of 100 heuristic tree searches, each search beginning with a distinct randomized maximum parsimony starting tree.

C.2.3 DNA extraction for real-time PCR

Genomic DNA was extracted from wild, unwashed colonies of *Trichodesmium* collected Aug 2003 in the North Pacific. Samples were stored at -20ºC in Delong Buffer (40 mM EDTA, 0.75M sucrose, 50 mM Tris-HCl, pH=8.3). In preparation for DNA extraction, Trition X-100 (final concentration of 5%) and lysozyme (10 mg/ml) were added to samples. Samples were then incubated at 37ºC for six hours, with occasional inverting of the tube. Proteinase K (final concentration of 1mg/mL) was added to samples which were then incubated at 70ºC for 40 min.

DNA extraction and purification was carried out using a standard phenol-chloroform procedure with ethanol precipitation. The purity of the extracted colony DNA was high, A_{260}/A_{280} of 1.94, but showed significant shearing when visualized by agarose gel electrophoresis.

C.2.4 Primer design and real-time PCR

Primer sets for real-time PCR were designed using Primer Express software version 3.0 (Applied Biosystems). The predicted specificity of the primers was examined by checking against the Ribosomal Database Project II (RDP-II) database (Cole et al., 2007). The primer sequences designed for this study are as follows: ST597F, 5'-GTG TAG GCG GAT ATA CAA GTT T-3'; ST703R, 5'-CAC TAG GAG TTC CAC ACT C-3'; Trich386F, 5'-GGA AGA AGA TCT GAC GGT ACC AA-3'; Trich485R, 5'-CCC AAT CAT TCC GGA TAA CG-3'. Real-time PCR was performed using the 7300 Real-Time PCR System and Sequence Detection Software version 1.3.1 (Applied Biosystems). Extracted *Trichodesmium* colony DNA was serially diluted (1:10), and each dilution was used as a template to ensure consistent results. Three independent master mixes were prepared for each template dilution and primer set, and each master mix was tested in duplicate. Standard DNA and negative-control reactions were also performed, both tested in duplicate. Reactions were performed in a volume of 25 μL, using Brilliant SYBR Green QPCR Master Mix (Stratagene) and either the *Trichodesmium* spp.-specific set of primers or the TAA-specific set of primers. All primers were used at a concentration of 0.2 μM. An initial denaturing step at 95ºC for 10 min was followed by 40 cycles of 95ºC for 30 s, 60ºC

for 45 s, and 72ºC for 45 s. A melt curve profile of the PCR products, run immediately after amplification and carried out in a range of 60ºC to 95ºC, confirmed the lack of nonspecific PCR products and primer dimers. Real-time PCR data was analyzed using Microsoft Excel, version 10 for Mac.

C.3 Results

C.3.1 Diversity of *Beta* and *Gammaproteobacteria*

Gammaproteobacteria associated with *Trichodesmium* included *Vibrio/Photobacterium*, *Pseudomonas*, *Acinetobacter*, *Psychrobacter*, *Alteromonas/Haliea* and *Pseudoalteromonas* as well as two difficult to classify sequences, EN530P15 and EN530P26 (Figure C.1). Some sequences grouped closely with other sequences found associated with a marine-host: EN530T53 and EN530Spnt5 grouped with a coral-associated clone, E8Altermonas grouped with a dinoflagellate-associated isolate, Sar530Deep20 grouped with alga and marine particle associated *Psychrobacter* isolates, and EN530P26 grouped with a sponge-associated clone.

All *Betaproteobacteria* sequences were found in one OTU, SAR27TDS4, which has a nearest-neighbor sequence in a metagenomic database from the Hawaii Ocean Time series ALOHA station.

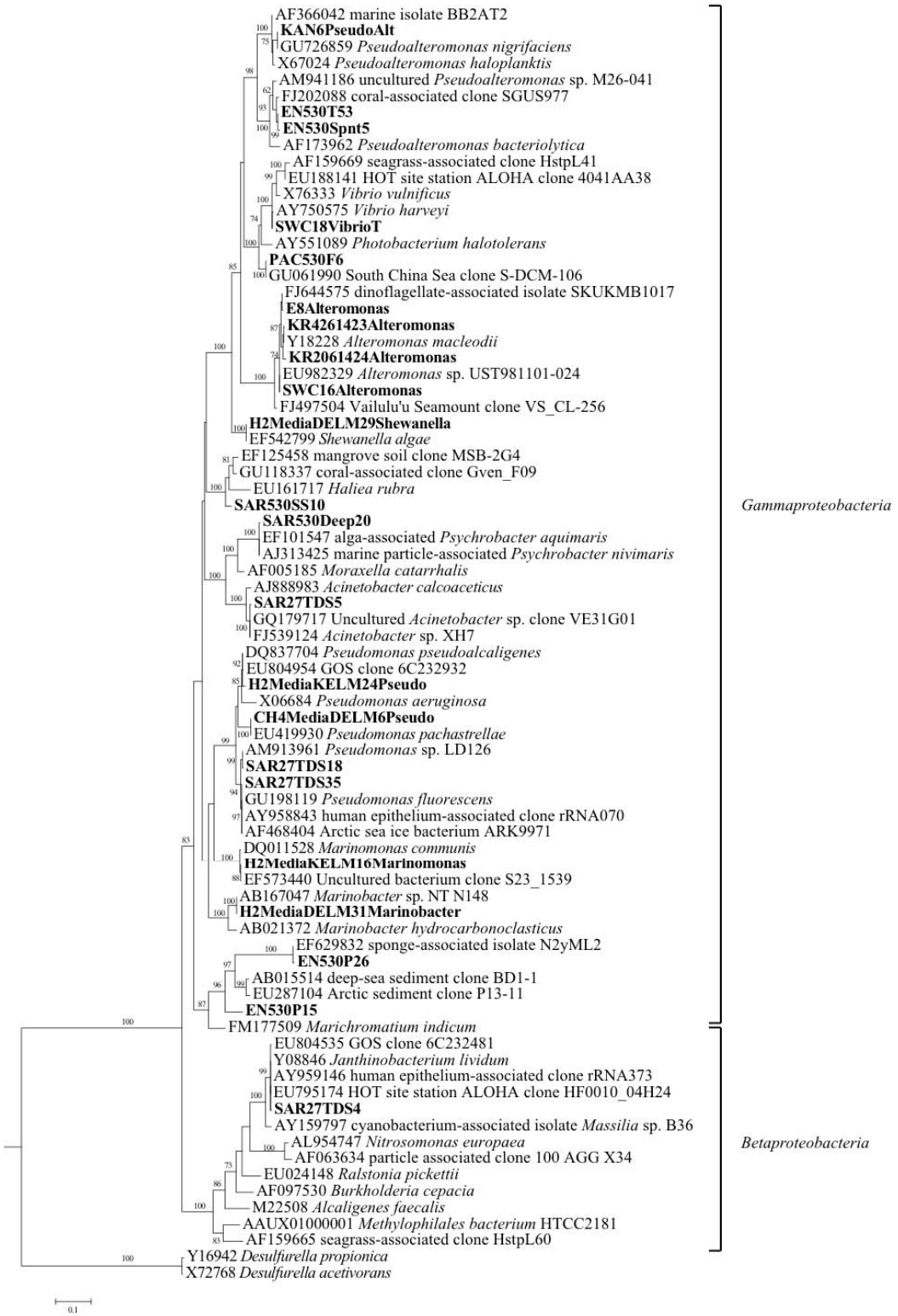

Figure C.1. Maximum likelihood tree of partial 16S rRNA gene sequences from *Trichodesmium*-associated *Beta* and *Gammaproteobacteria* and select reference sequences. Sequences associated with *Trichodesmium* are shown in bold. Boostrap values over 60% are shown and are based on 100 replicates. Scale bar represents 0.1 nucleotide substitutions per site.

C.3.2 Abundance of TaA in *Trichodesmium* colonies

The relative abundance of 16S rRNA gene copy numbers of TaA and *Trichodesmium* were compared within the same colonies using an absolute quantitative real-time PCR assay. Standard curves of threshold cycle number as a function of DNA concentration were generated from plasmid DNA containing a fragment of either *Trichodesmium* or the TaA representative PAC530F63 16S rRNA gene. Standard curves bracketed the cycle numbers observed in the colony assays. The slope of the *Trichodesmium* plasmid standard curve was –3.552, the correlation coefficient was >0.98, and the PCR amplification efficiency was 1.912, according to the equation $E = 10^{(-1/\text{slope})}$ (Rasmussen, 2001). The slope of the TAA plasmid standard curve was –3.324, the correlation coefficient was >0.99, and the PCR amplification efficiency was 1.999.

Assays of colony DNA with *Trichodesmium*- and TaA-specific 16S rRNA primers were compared with their respective standard curves in order to quantify the relative amount of TaA DNA in *Trichodesmium* colonies. TaA 16S rRNA gene copy numbers were 23-28 times more abundant than *Trichodesmium* 16S rRNA gene copy numbers. This abundance was calculated as the ratio of the average TaA 16S rRNA gene copy numbers obtained with TaA-specific primers and the average *Trichodesmium* 16S rRNA gene copy numbers obtained with *Trichodesmium*-specific primers, both using the same extracted colony genomic DNA as a template. One nanogram of colony DNA contained approximately 3.57×10^3 copies of *Trichodesmium* 16S rRNA gene and 1.01×10^5 copies of TaA 16S rRNA gene.

C.4 Discussion

Analysis of 16S rRNA sequences associated with geographically diverse *Trichodesmium* colonies shows a relatively diverse population of *Gammaproteobacteria* and relatively few *Betaproteobacteria*. The goal of this work was to discover potentially specific relationships between *Trichodesmium* and its associated microbial community. The finding of sequences both in the surrounding seawater and in *Trichodesmium* colonies would suggest that the sequences recovered from *Trichodesmium* colonies are just a subset of free-living bacteria. Differences between sequences recovered from *Trichodesmium* colonies to those found in the surrounding seawater can be detected by examining the results of the Global Ocean Survey (GOS), which collected over 6000 16S rRNA gene sequences from seawater samples (Shaw et al., 2008). Filtration of samples excluded *Trichodesmium* from the GOS dataset. Comparison of sequences from *Trichodesmium* colonies to GOS data shows relatively few sequences with high similarity to GOS sequences, with the exception of SAR27TDS4 and H2MediaKELM24Pseudo (Figure C.1). This suggests a population associated with *Trichodesmium* that is relatively distinct from that found in the surrounding seawater.

A recent metatranscriptomic study of *Trichodesmium* colonies also finds that the *Trichodesmium*-associated community is different from free-living bacterioplankton communities (Hewson et al., 2009). Comparison of sequences retrieved in the current study with the dataset obtained in the metatransciptomic dataset shows little overlap in the Beta and Gammaproteobacterial communities, with

two sequences, SAR530SS10 and EN530T53, showing >99% identity with sequences in the metatranscriptomic dataset. There are several factors that may explain the differences found between the current study's dataset and the metatransciptomic dataset. Firstly, current study is a DNA analysis of community diversity, while the latter is a gene expression analysis of the community. Therefore, more metabolically active microbes may dominate sequences in the latter dataset. Also, the metabolic status of colonies collected in both studies may not be comparable. Colonies in the metatransciptomic study were gathered during a visible, high-biomass surface *Trichodesmium* bloom (Hewson et al., 2009), while these conditions were not present when surface colonies were collected in the current study (pers. comm. Elizabeth Mann). Additionally, the handling of colonies at the time of collection differed in the studies. Colonies were rinsed with filtered seawater before processing in the current study, while colonies were not rinsed in the metatranscriptomic study.

Real-time qPCR was used in order to determine the abundance of the unique TaA clade of bacteria within the *Trichodesmium*-associated community. Our results show that TaA SSU rRNA gene copy numbers are more than 20-fold greater than those of *Trichodesmium* in a wild colony. However, converting these gene copy numbers to cell abundance is complicated by the fact that while *Trichodesmium erythraeum* has two copies of the 16S rRNA genes, the number of 16S rRNA gene copies and therefore the exact cell ratio of TaA to *Trichodesmium* is unknown. Nevertheless, if TaA relies on *Trichodesmium* colonies to provide a nutrient-rich habitat, a rRNA gene copy number of between one and three is a reasonable estimate since symbionts and parasites inhabiting relatively stable environments typically have

low copy numbers of rRNA operons (Klappenbach et al., 2000). This would yield a cell ratio of TaA to *Trichodesmium* as high as 56 or as low as 15 to 16. Even if TaA is a rapidly growing opportunist like *E. coli*, with seven rRNA operon copies, the ratio would still be six or more TaA cells per *Trichodesmium* cell. Although there are uncertainties involved in these estimates, the results do support the conclusion that TaA is an abundant member of the *Trichodesmium*-associated community.

Examination of the metatransciptomic dataset reveals the opposite trend, however, where the abundance of *Trichodesmium* 16S rRNA is orders of magnitude higher than that of TaA. While this may indicate that TaA is less metabolically active than *Trichodesmium*, a more rigorous study comparing gene expression and gene abundance in the same sample needs to be carried out before a firm conclusion can be reached. Nevertheless, the finding of TaA in the metatransciptomic dataset indicates that this unique group of bacteria is metabolically active within *Trichodesmium* colonies.

Appendix D. The diversity and localization of archaea in the tropical sponge *Corticium* sp.

D.1 Introduction

Marine sponges (phylum Porifera) are significant members of the benthic community is terms of biomass and influence on biological processes (Taylor et al., 2007). They are sessile, filter-feeding organisms, filtering dissolved organic matter, particulates, and microorganisms from water through an extensive canal system. Water enters through pores on the outer surface, ostia (Figure D.1). Specialized flagellated cells, called choanocytes, form a series of chambers, called the choanoderm, and function to pump water through the sponge and filter out food particles. After filtration, water flows out through the osculum. Food particles are transferred to the mesohyl, which is an extensive layer of connective tissue. The mesohyl is also the location of a dense community of microorganisms, as well as the location of embryo brooding. Embryos are brooded within the sponge, and then fully developed larvae are released into the water column.

Sponges harbor diverse and abundant microbial communities (Taylor et al., 2007). The precise roles of the microorganisms in the sponge is not entirely clear, but some possible roles are to provide food, protect from ultraviolet light, and chemically defend against predators (Holmes and Blanch, 2006). Considerable interest has been paid to sponges because they are a prolific source of novel bioactive compounds (Faulkner, 2001). However, more recently, sponges have gained attention for the discovery of their specific associations with archaea, a group of microbes once thought to predominately inhabit extreme environments. This specific association

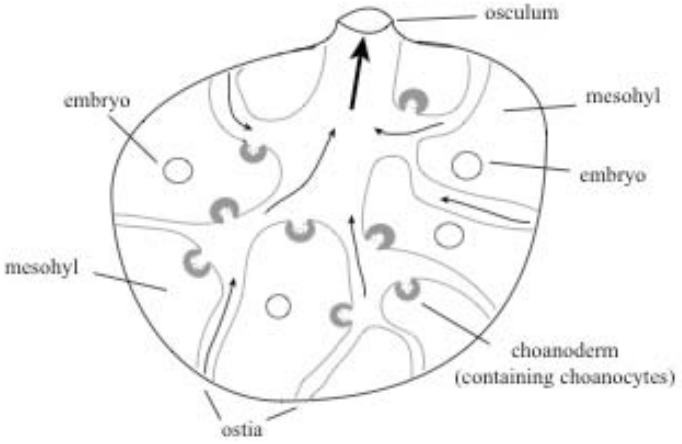

Figure D.1. Schematic representation of a cross-section of a marine sponge. Arrows indicate the flow of water through the sponge.

was first reported between the sponge *Axinella mexicana* and its associated archaeon, *Cenarchaeum symbiosum* (Preston et al., 1996). Since this initial discovery, numerous reports of sponge-associated archaea have appeared (Webster et al., 2001; Margot et al., 2002; Lee et al., 2003; Holmes and Blanch, 2006). The roles of these archaea were initially unknown. Then with the report of mesophilic crenarchaeota capable of ammonia oxidation (Treusch et al., 2005) combined with the finding of ammonia monooxygenase (*amo*) gene homologs in the genome of *C. symbiosum* (Hallam et al., 2006a; Hallam et al., 2006b), it was suggested that these sponge-associated archaea may be involved in the cycling of nitrogen within the sponge host.

Previous work with the tropical sponge *Corticium* sp. revealed the vertical transmission of both bacteria and archaea by localizing the microbial community in developing embryos within the adult brood chambers (Sharp et al., 2006). While the bacterial community was characterized using 16S rRNA gene sequences, the archaeal community was not. This study presents the characterization of the archaeal

community associated with the sponge *Corticium* sp. Specific probes were designed and used to localize the archaeal community in embryos within host tissue, suggesting the vertical transmission of these archaea. Additionally, potential functions of the archaeal community are proposed through the detection of *amoA* genes.

D.2 Methods

D.2.1 Sample collection

Corticium sp. samples were collected by scuba diving in the Republic of Palau at depths ranging from 5 to 20 m in the May 2007. Samples were rinsed with filtered seawater three times before processing. For DNA, samples were preserved in RNAlater (20 mM EDTA, 25 mM sodium citrate, 700 g/L ammonium sulfate, pH 5.2) and stored at -20°C until use. For FISH, samples were fixed in 4% paraformaldehyde in phosphate buffer (15 mM Na_2HPO_4, 5 mM NaH_2PO_4, 0.5 M NaCl, pH 7.4) for 2 hrs at room temperature, then transferred to 70% ethanol and stored at -20°C until use. The sample used for the current study, 07-085, was collected from the site Big Dropoff in Palau.

D.2.2 Archaeal 16S rRNA gene clone library

Approximately 10 mg of sponge tissue was extracted for DNA. Sponge sample was sliced into small pieces with a sterile razor blade and then homogenized with a mortar and pestle in lysis buffer (20 mg/ml lysozyme, 20 mM Tris-Cl pH 8.0, 2 mM EDTA, 1.2% Triton X-100). This homogenate was incubated at 37°C

overnight. In the morning, Proteinase K (final concentration of 1mg/mL) was added to samples, which were then incubated at 70ºC for 30 min. DNA extraction and purification was carried out using a standard phenol-chloroform procedure with ethanol precipitation. PCR with general archaeal 16S rRNA gene primers was performed with the primers Arch21F and Arch958R (Table D.1) at the following conditions: an initial denaturation step for 2 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 70°C, and then a final extension step for 7 min at 72°C. PCR products were analyzed by gel electrophoresis. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen). Purified PCR products were cloned with the TOPO TA Cloning Kit (Invitrogen) with the pCR4 vector. Forty-one clones were chosen for plasmid extraction and sequencing. Purified plasmids with archaeal 16S rRNA gene amplicons were sequenced with M13F/R primers.

D.2.3 Phylogenetic analysis

Chimeras were checked for using Bellerophon (Huber et al., 2004). Diversity assessment and operational taxonomic unit (OTU) assignments were performed by the software package DOTUR (Schloss and Handelsman, 2005), using a 98% cutoff. *Corticium*-associated archaeal sequences were aligned using SILVA's SINA webaligner (Pruesse et al., 2007). Aligned near-neighbors and reference sequences were downloaded from SILVA's rRNA database. The multiple sequence alignment was edited in Bioedit (Hall, 1999). Phylogenetic reconstructions were performed using RAxML v. 7.0.4 (Stamatakis, 2006) using the General Time Reversible model

of nucleotide substitution under the Γ model of rate heterogeneity (GTRGAMMA) with 100 bootstrap replicates. The selected tree topology had the highest likelihood score out of 100 heuristic tree searches, each search beginning with a distinct randomized maximum parsimony starting tree.

D.2.4 Fluorescence in situ hybridization (FISH)

Specific probes were designed for the *Corticium* sp.-associated archaea sequences produced in the clone library analysis using the ARB software package (Ludwig et al., 2004). Probe specificity was checked against the Ribosomal Database Project II (RPD-II) database using the Probe Match function (Cole et al., 2007). Successful hybridization of specific-archaea probes required the use of unlabeled helper probes as described by Fuchs et al. (2000). These unlabeled probes bind adjacent to the fluorescent-probe binding site, allowing better access to the target site by "opening" the rRNA molecule.

Samples were embedded in paraffin wax and cross-sectioned to 10 μm. Sections were deparaffinized with a 5 min treatment in xylene, followed by a 5 min treatment in ethanol, and then rinsed in Milli-Q water. Slides were air-dried. Sections were hybridized with 5 ng/μL probe (final concentration) in hybridization buffer (35% formamide, 0.9 M NaCl, 20 mM Tris-HCl [pH 7.4], 0.01% sodium dodecyl sulfate) for 2 hrs at 46°C. In the case of the specific archaeal probes, helper probes were also used at a final concentration of 5 ng/μL. Mismatch probes were also used for negative controls. Sequences of all probes used in this study are listed in Table D.1. After hybridization, the slides were incubated at 48°C in wash buffer (0.7

M NaCl, 20 mM Tris-HCl [pH 7.4], 50 mM EDTA, 0.01% sodium dodecyl sulfate) for 20 min. The wash buffer was rinsed off with Milli-Q water, slides were air dried, and then mounted in a 4:1 Citifluor (Citifluor Ltd.):VectaShield (Vector Labs) mounting medium. Slides were visualized on a Zeiss AxioImager laser scanning confocal microscope with LSM 5 Pascal Version 4.0 imaging software.

D.2.5 RT-PCR of *amoA* genes

The presence of both bacterial and archaeal *amoA* genes was checked by PCR with DNA from *Corticium* sp. sample 07-085 using separate sets of bacterial-specific and archaeal-specific *amoA* gene primers (Table D.1). Amplicons were sequenced to ensure the presence of the expected product. RNA from *Corticium* sample 07-085 was extracted using the Tri Reagent (Sigma) according to manufacturer's instructions. The RNA was then further purified using the illustra RNAspin Mini Kit (GE). Reverse-transcriptase PCR (RT-PCR) was carried out using the SuperScript III One-step RT-PCR System with Platinum *Taq* DNA Polymerase (Invitrogen) with the following PCR cycling conditions: a cDNA synthesis step for 30 min at 50°C, followed by a denaturation step for 2 min at 94°C, followed by 30 cycles consisting of 1 min at 94°C, 1 min at 55°C, 1 min at 72°C, and then a final extensions step for 7 min at 72°C. RT-PCR assays included negative controls for DNA and RNA contamination. To check for the absence of contaminating genomic DNA in the RNA preparation, a control PCR containing no reverse transcriptase enzyme was carried out with the RNA preparation using the same primers and cycling conditions.

Table D.1. List of oligonucleotide primers and probes used in this study.

Probe/primer	Sequence (5'-3')	Reference
Arch21F	TTCCGGTTGATCCYGCCGGA	DeLong, 1992
Arch958R	YCCGGCGTTGAMTCCAATT	Delong, 1992
Arch915	GTGCTCCCCCGCCAATTCCT	Stahl and Amann, 1991
Arch915NON	GTGCTACCCCGCCAATTCCT	Sharp et al., 2007
EUB338	GCTGCCTCCCGTAGGAGT	Amann et al., 1990
EUBNON	ACTCCTACGGGAGGCAGC	Wallner et al., 1993
CArch758	GCATGGTTTACAGCTGGGA	This study
CArch758NON	GCATGGTCTACAGCTGGGA	This study
CA758help740	CTACCCGGGTATCTAATC	This study
CA758help777	CAARGCATCACTGAGTTT	This study
Arch-amoAF	STAATGGTCTGGCTTAGACG	Francis et al., 2005
Arch-amoAR	GCGGCCATCCATCTGTATGT	Francis et al., 2005
Bac-amoA-1F	GGGGTTTCTACTGGTGGT	Rotthauwe et al., 1997
Bac-amoA-2R	CCCCTCKGSAAAGCCTTCTTC	Rotthauwe et al., 1997

D.3 Results

D.3.1 Phylogenetic analysis of archaea associated with *Corticium* sp.

The archaeal community of *Corticium* sp. was characterized by constructing a clone library of archaeal 16S rRNA gene sequences amplified from whole adult tissue. Forty-one clones were sequenced, returning 20 unique sequences. Operational taxonomic units (OTU) were determined by the software package DOTUR, defining 2 distinct OTUs, *Corticium* associated clone 01 and *Corticium* associated clone 03. Comparison to the GenBank (NCBI) database shows that clone 01 has 99% identity to an uncultivated marine crenarchaeote, while clone 03 has 97% identity to the ammonia-oxidizing crenarchaeote *Nitrosopumilus maritimus.* These results suggest that the archaeal community of *Corticium* sp. is relatively homogenous, compared to the diverse group of bacteria found within this species (Sharp et al., 2007).

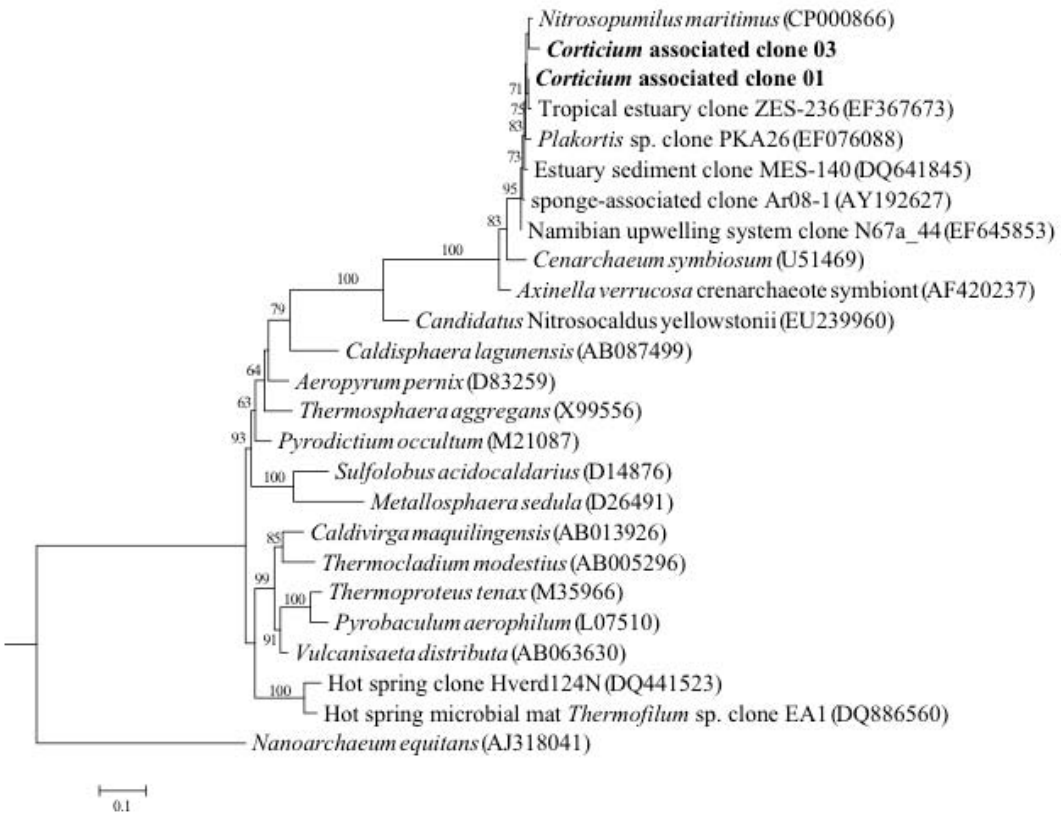

Figure D.1. Maximum likelihood tree of partial archaeal 16S rRNA gene sequences associated with the marine sponge *Corticium* sp. and select reference sequences. Sequences associated with *Corticium* sp. are shown in bold. Boostrap values over 60% are shown and are based on 100 replicates. Scale bar represents 0.1 nucleotide substitutions per site.

D.3.2 Localization of the archaeal community

Corticium sp. sections were hybridized with a CY5-labeled general eubacterial probe (EUB338) and a CY3-labeled general archaeal probe (Arch915) by FISH. The results were similar to those reported by Sharp et al. (2007). Clusters of bacteria and archaea are found around the periphery of early-stage embryos (Figure D.2B). In later developmental stage embryos, bacteria and archaea are found in the central cavity (Figure D.2C). In both cases, bacteria far outnumber archaea and are seen interspersed within bacterial clusters.

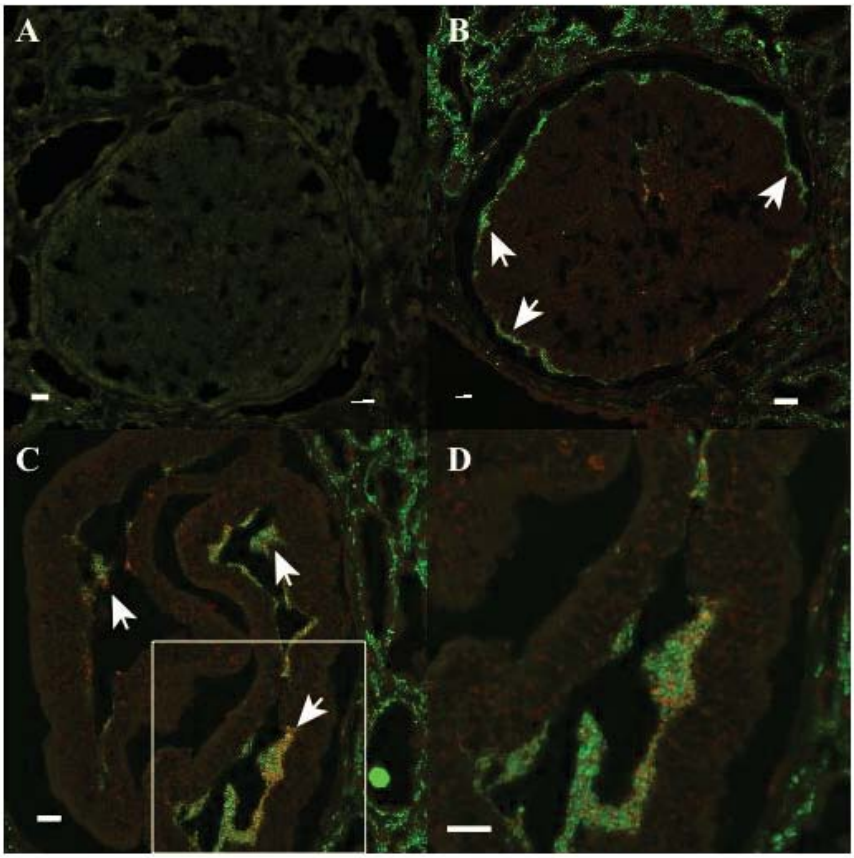

Figure D.2. FISH images of bacteria labeled with CY3-EUB338 (green) and archaea labeled with CY5-Arch915 (red) in developing *Corticium* sp. embryos. (A) Negative control of an early stage embryo with probes CY3-EUBNON and CY5-Arch915NON showing no hybridization. (B) Early-stage embryo showing clusters of bacteria and archaea (arrows) around the periphery of the developing embryo. Archaea are severely outnumbered by bacteria, therefore, the archaeal signal is somewhat drowned out in this image. The surrounding mesohyl, meanwhile, is densely packed with bacteria and some archaea. (C) Later developmental stage embryo, showing the localization of bacteria and archaea in the central cavities of the embryo (arrows). (D) Close up image of the area boxed in C. The archaea are seen interspersed within the clusters of bacteria. Scale bars in each image are 10 μm.

A specific oligonucleotide probe was designed to target both archaeal OTU sequences retrieved in the clone library analysis, CY3-CArch758. Unlabeled helper oligonucleotide probes (Fuchs et al., 2000) were required for successful hybridization. A one-base-mismatch probe was also designed and used to confirm specificity. Hybridization with CY5-Arch915 and CY3-CArch758 resulted in complete colocalization of both probe signals (Figure D.3), indicating that the

archaeal community in *Corticium* sp. is well represented in the clone library analysis. Localization of the specific probe in developing embryos supports the proposed vertical transmission of this community of archaea.

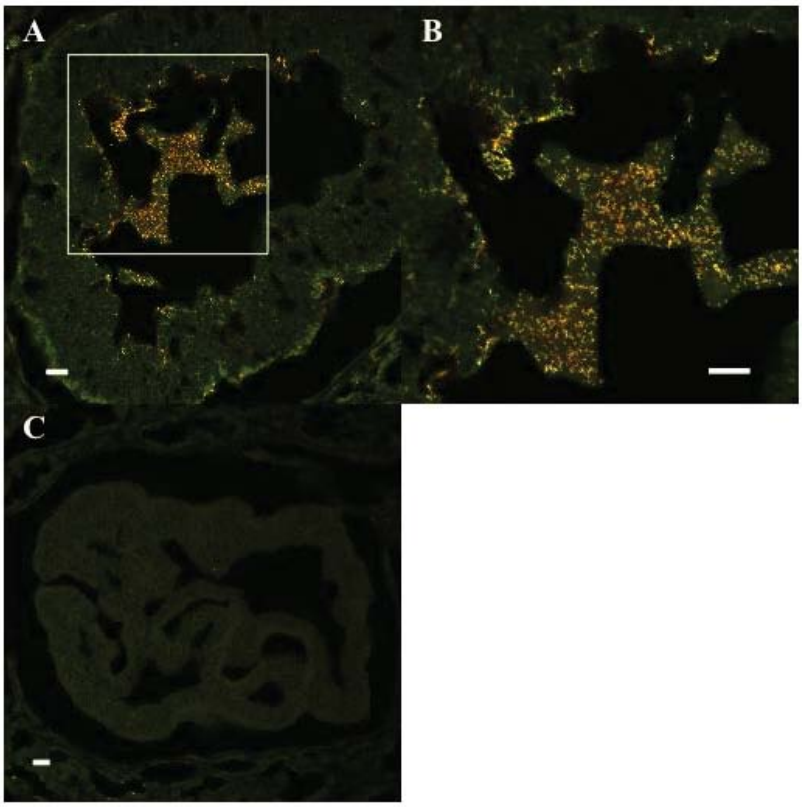

Figure D.3. FISH images of CY5-Arch915 (red) and CY3-CArch758 (green) dual-labeled archaea in a late developmental stage embryo. Colocalization of both probes results in a yellow signal. (A) Colocalization of the general (Arch915) and specific (CArch758) probes in archaea located in the central cavity of a late developmental stage embryo. (B) Close up image of the area boxed in A. (C) Negative control image with probes CY5-Arch915NON and CY3-CArch758NON showing no hybridization. Scale bars in each image are 10 µm.

D.3.3 Expression of archaeal *amoA* gene

To determine the potential role of *Corticium* sp.-associated archaea, the presence of archaeal *amoA* genes was examined. PCR of *Corticium* sp. DNA with general archaeal *amoA* gene primers confirmed the presence of these genes in the

sponge. PCR with general bacterial *amoA* gene primers did not produce a product (data not shown). To determine if the archaeal *amoA* genes are actively expressed within the host, reverse-transcriptase (RT) PCR was carried out. RNA was extracted from whole adult tissue *Corticium* sp. A cDNA library was constructed and archaeal *amoA* genes were amplified, indicating the active expression of these genes within the sponge (Figure D.4).

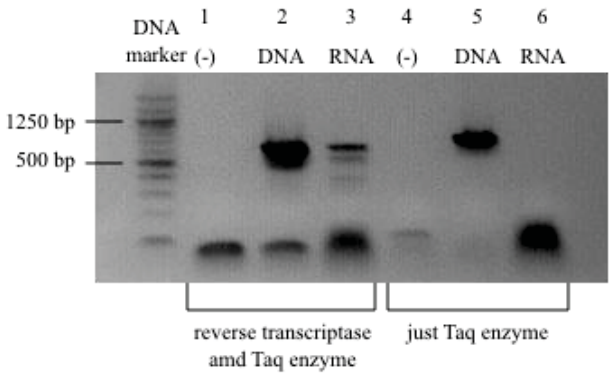

Figure D.4. RT-PCR assay of *Corticium* sp. RNA and controls. Lanes 1 and 4 show no template negative controls. Lanes 2 and 5 show positive amplification of DNA from *Corticium* sp. with archaeal *amoA* primers. Lane 3 shows the positive amplification of archaeal *amoA* genes after cDNA synthesis. Lane 6 is a negative control showing the absence of DNA contamination in the RNA template.

D.4 Discussion

Vertical transmission of associated microbes suggests a specific relationship between the microbes and the host. In this work, the archaeal community associated with the marine sponge *Corticium* sp. was characterized and evidence was shown for the vertical transmission of this archaeal community. In contrast to the diverse group of bacteria detected in this sponge (Sharp et al., 2007), the archaeal community is a closely related group. Localization of archaea to developing embryos supports

previous localization studies (Sharp et al., 2007). The vertical transmission of this community suggests a specific relationship with the host.

The potential role of these archaea was explored through a RT-PCR assay. Archaeal *amoA* gene transcripts were detected in *Corticium* sp., suggesting the active oxidation of ammonia in the sponge. The association of ammonia-oxidizing archaea with marine sponges has been reported previously (Hallam et al., 2006b; Steger et al., 2008). However, these reports have not examined the expression of this gene in the sponge. The accumulation of metabolic wastes such as ammonia may be detrimental to the sponge. Therefore, the association with ammonia oxidizers may be beneficial to the sponge. Bacterial *amoA* genes were not detected in this study. This suggests that archaea, not bacteria, are the main contributors to nitrification in *Corticium* sp., which has been shown to be the case in other environments (Leininger et al., 2006; Wuchter et al., 2006).

References

Acinas, Silvia G, Vanja Klepac-Ceraj, Dana E Hunt, Chanathip Pharino, Ivica Ceraj, Daniel L Distel, and Martin F Polz. 2004. "Fine-scale phylogenetic architecture of a complex bacterial community." *Nature* 430 (6999) (July 29): 551-554. doi:10.1038/nature02649.

Actis, L A, W Fish, J H Crosa, K Kellerman, S R Ellenberger, F M Hauser, and J Sanders-Loehr. 1986. "Characterization of anguibactin, a novel siderophore from Vibrio anguillarum 775(pJM1)." *Journal of Bacteriology* 167 (1) (July): 57-65.

Amann, R I, B J Binder, R J Olson, S W Chisholm, R Devereux, and D A Stahl. 1990. "Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations." *Applied and Environmental Microbiology* 56 (6) (June): 1919-1925.

Anand, Swadha, M V R Prasad, Gitanjali Yadav, Narendra Kumar, Jyoti Shehara, Md Zeeshan Ansari, and Debasisa Mohanty. 2010. "SBSPKS: structure based sequence analysis of polyketide synthases." *Nucleic Acids Research* 38 (Web Server issue) (July): W487-496. doi:10.1093/nar/gkq340.

Andrews, Simon C, Andrea K Robinson, and Francisco Rodríguez-Quiñones. 2003. "Bacterial iron homeostasis." *FEMS Microbiology Reviews* 27 (2-3) (June): 215-237.

Ansari, Mohd Zeeshan, Gitanjali Yadav, Rajesh S Gokhale, and Debasisa Mohanty. 2004. "NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases." *Nucleic Acids Research* 32 (Web Server issue) (July 1): W405-413. doi:10.1093/nar/gkh359.

Arnow, L. Earle. 1937. "Colorimetric determination of the components of 3,4-dihydroxyphenyalaninetyrosine mixtures." *Journal of Biological Chemistry* 118 (2) (April 1): 531 -537.

Bachhawat, A. K., and S. Ghosh. 1987. "Iron Transport in Azospirillum brasilense: Role of the Siderophore Spirilobactin." *Microbiology* 133 (July 1): 1759-1765. doi:10.1099/00221287-133-7-1759.

Bachmann, Brian O, and Jacques Ravel. 2009. "Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data." *Methods in Enzymology* 458: 181-217. doi:10.1016/S0076-6879(09)04808-3.

Baichoo, N., and J. D. Helmann. 2002. "Recognition of DNA by Fur: a Reinterpretation of the Fur Box Consensus Sequence." *Journal of Bacteriology* 184 (November 1): 5826-5832. doi:10.1128/JB.184.21.5826-5832.2002.

Baig, B H, I K Wachsmuth, and G K Morris. 1986. "Utilization of exogenous siderophores by Campylobacter species." *Journal of Clinical Microbiology* 23 (3) (March): 431-433.

Balibar, Carl J., Frédéric H. Vaillancourt, and Christopher T. Walsh. 2005. "Generation of D Amino Acid Residues in Assembly of Arthrofactin by Dual Condensation/Epimerization Domains." *Chemistry & Biology* 12 (11) (November): 1189-1200. doi:10.1016/j.chembiol.2005.08.010.

Barbeau, K, E L Rue, K W Bruland, and A Butler. 2001. "Photochemical cycling of iron in the surface ocean mediated by microbial iron(III)-binding ligands." *Nature* 413 (6854) (September 27): 409-413. doi:10.1038/35096545.

Barbeau, Katherine, Eden L. Rue, Charles G. Trick, Kenneth W. Bruland, and Alison Butler. 2003. "Photochemical reactivity of siderophores produced by marine heterotrophic bacteria and cyanobacteria, based on characteristic Fe(III) binding groups." *Limnology and Oceanography* 48: 1069-1078. doi:10.4319/lo.2003.48.3.1069.

Barbeau, Katherine, Guangping Zhang, David H Live, and Alison Butler. 2002. "Petrobactin, a photoreactive siderophore produced by the oil-degrading marine bacterium Marinobacter hydrocarbonoclasticus." *Journal of the American Chemical Society* 124 (3) (January 23): 378-379.

Bellenger, J. P., T. Wichard, A. B. Kustka, and A. M. L. Kraepiel. 2008. "Uptake of molybdenum and vanadium by a nitrogen-fixing soil bacterium using siderophores." *Nature Geoscience* 1 (4) (March): 243-246. doi:10.1038/ngeo161.

Bernhard Schwyn, and J.B. Neiland. 1987. "Universal chemical assay for the detection and determination of siderophores." *Analytical Biochemistry* 160: 47-56.

Betcher, Meghan. 2011. Insights into the shipworm (Bivalvia: Teredinidae)-bacterial symbiosis: Novel isolates and micrographic localization studies. M.S. Thesis, Oregon Health & Science University.

Bigliardi, Elisa, Maria Gloria Selmi, Silvia Corona, Claudio A. Bandi, and Luciano Sacchi. 1995. "Membrane systems in endocytobiosis III. Ultrastructural features of symbionts and vacuolar membrane in bacteriocytes of the wood-eating cockroach Cryptocercus punctulatus (Dictyoptera, Cryptocercidae)." *Bolletino di zoologia* 62 (January): 235-238. doi:10.1080/11250009509356070.

Bishop, P E, and R D Joerger. 1990. "Genetics and Molecular Biology of Alternative Nitrogen Fixation Systems." *Annual Review of Plant Physiology and Plant Molecular Biology* 41 (June): 109-125. doi:10.1146/annurev.pp.41.060190.000545.

Bister, Bojan, Daniel Bischoff, Graeme J. Nicholson, Marianne Valdebenito, Kathrin Schneider, Günther Winkelmann, Klaus Hantke, and Roderich D. Süssmuth. 2004. "The structure of salmochelins: C-glucosylated enterobactins of Salmonella enterica." *BioMetals* 17 (August): 471-481. doi:10.1023/B:BIOM.0000029432.69418.6a.

Breznak, J A, and A Brune. 1994. "Role of Microorganisms in the Digestion of Lignocellulose by Termites." *Annual Review of Entomology* 39 (1) (January): 453-487. doi:10.1146/annurev.en.39.010194.002321.

Brickman, T J, and M A McIntosh. 1992. "Overexpression and purification of ferric enterobactin esterase from Escherichia coli. Demonstration of enzymatic hydrolysis of enterobactin and its iron complex." *The Journal of Biological Chemistry* 267 (17) (June 15): 12350-12355.

Bruner, Steven D, Thomas Weber, Rahul M Kohli, Dirk Schwarzer, Mohamed A Marahiel, Christopher T Walsh, and Milton T Stubbs. 2002. "Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE." *Structure* 10 (March): 301-310. doi:10.1016/S0969-2126(02)00716-5.

Butler, Alison. 2005. "Marine Siderophores and Microbial Iron Mobilization." *BioMetals* 18 (August): 369-374. doi:10.1007/s10534-005-3711-0.

Challis, Gregory L, Jacques Ravel, and Craig A Townsend. 2000. "Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains." *Chemistry & Biology* 7 (3) (March): 211-224. doi:10.1016/S1074-5521(00)00091-0.

Challis, Gregory L. 2005. "A Widely Distributed Bacterial Pathway for Siderophore Biosynthesis Independent of Nonribosomal Peptide Synthetases." *ChemBioChem* 6 (April 8): 601-611. doi:10.1002/cbic.200400283.

Choi, Dong-Sic, Dae-Kyum Kim, Seng Jin Choi, Jaewook Lee, Jun-Pyo Choi, Sangchul Rho, Sang-Hyun Park, Yoon-Keun Kim, Daehee Hwang, and Yong Song Gho. 2011. "Proteomic analysis of outer membrane vesicles derived from Pseudomonas aeruginosa." *Proteomics* 11 (August): 3424-3429. doi:10.1002/pmic.201000212.

Chu, Byron C., Alicia Garcia-Herrero, Ted H. Johanson, Karla D. Krewulak, Cheryl K. Lau, R. Sean Peacock, Zoya Slavinskaya, and Hans J. Vogel. 2010. "Siderophore uptake in bacteria and the battle for iron with the host; a bird's eye view." *BioMetals* 23 (4) (July): 601-611. doi:10.1007/s10534-010-9361-x.

Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje. 2007. "The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data." *Nucleic Acids Research* 35 (January 3): D169-D172. doi:10.1093/nar/gkl889.

Conti, E, T Stachelhaus, M A Marahiel, and P Brick. 1997. "Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S." *The EMBO Journal* 16 (14) (July 16): 4174-4183.

Cornelis, Pierre, and Sandra Matthijs. 2002. "Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines." *Environmental Microbiology* 4 (12) (December): 787-798. doi:10.1046/j.1462-2920.2002.00369.x.

Crosa, J H, L L Hodges, and M H Schiewe. 1980. "Curing of a plasmid is correlated with an attenuation of virulence in the marine fish pathogen Vibrio anguillarum." *Infection and Immunity* 27 (3) (March): 897-902.

Crosa, Jorge H. 1980. "A plasmid associated with virulence in the marine fish pathogen Vibrio anguillarum specifies an iron-sequestering system." *Nature* 284 (5756) (April): 566-568. doi:10.1038/284566a0.

D'Onofrio, Anthony, Jason M. Crawford, Eric J. Stewart, Kathrin Witt, Ekaterina Gavrish, Slava Epstein, Jon Clardy, and Kim Lewis. 2010. "Siderophores from Neighboring Organisms Promote the Growth of Uncultured Bacteria." *Chemistry & Biology* 17 (3) (March): 254-264. doi:10.1016/j.chembiol.2010.02.010.

DeLong, E F. 1992. “Archaea in coastal marine environments.” *Proceedings of the National Academy of Sciences of the United States of America* 89 (12) (June 15): 5685-5689.

Dertz, Emily A., Jide Xu, Alain Stintzi, and Kenneth N. Raymond. 2006. “Bacillibactin-Mediated Iron Transport in Bacillus subtilis.” *Journal of the American Chemical Society* 128 (January): 22-23. doi:10.1021/ja055898c.

Distel, D L, and C M Cavanaugh. 1994. “Independent phylogenetic origins of methanotrophic and chemoautotrophic bacterial endosymbioses in marine bivalves.” *Journal of Bacteriology* 176 (7) (April): 1932-1938.

Distel, D L, E F DeLong, and J B Waterbury. 1991. “Phylogenetic characterization and in situ localization of the bacterial symbiont of shipworms (Teredinidae: Bivalvia) by using 16S rRNA sequence analysis and oligodeoxynucleotide probe hybridization.” *Applied and Environmental Microbiology* 57 (8) (August): 2376-2382.

Distel, D L, D J Lane, G J Olsen, S J Giovannoni, B Pace, N R Pace, D A Stahl, and H Felbeck. 1988. “Sulfur-oxidizing bacterial endosymbionts: analysis of phylogeny and specificity by 16S rRNA sequences.” *Journal of Bacteriology* 170 (6) (June): 2506-2510.

Distel, D L, H K Lee, and C M Cavanaugh. 1995. “Intracellular coexistence of methano- and thioautotrophic bacteria in a hydrothermal vent mussel.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (21) (October 10): 9598-9602.

Distel, D. L., D. J. Beaudoin, and W. Morrill. 2002a. “Coexistence of Multiple Proteobacterial Endosymbionts in the Gills of the Wood-Boring Bivalve Lyrodus pedicellatus (Bivalvia: Teredinidae).” *Applied and Environmental Microbiology* 68 (12) (December): 6292-6299. doi:10.1128/AEM.68.12.6292-6299.2002.

Distel, Daniel L, Mehwish Amin, Adam Burgoyne, Eric Linton, Gustaf Mamangkey, Wendy Morrill, John Nove, Nicole Wood, and Joyce Yang. 2011. “Molecular phylogeny of Pholadoidea Lamarck, 1809 supports a single origin for xylotrophy (wood feeding) and xylotrophic bacterial endosymbiosis in Bivalvia.” *Molecular Phylogenetics and Evolution* 61 (2) (November): 245-254. doi:10.1016/j.ympev.2011.05.019.

Distel, Daniel L. 2003. The Biology of Marine Wood Boring Bivalves and Their Bacterial Endosymbionts. In *Wood Deterioration and Preservation*, ed. Barry Goodell, Darrel D. Nicholas, and Tor P. Schultz, 845:253-271. Washington, DC: American Chemical Society, March 31. http://pubs.acs.org/doi/abs/10.1021/bk-2003-0845.ch014.

Distel, Daniel L., Wendy Morrill, Noelle MacLaren-Toussaint, Dianna Franks, and John Waterbury. 2002b. “Teredinibacter turnerae gen. nov., sp. nov., a dinitrogen-fixing, cellulolytic, endosymbiotic gamma-proteobacterium isolated from the gills of wood-boring molluscs (Bivalvia: Teredinidae).” *International Journal of Systematic and Evolutionary Microbiology* 52 (6) (November): 2261-2269. doi:10.1099/ijs.0.02184-0.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2011. *Geneious v5.4, Available from http://www.geneious.com/*.

Dubilier, N, C Mülders, T Ferdelman, D de Beer, A Pernthaler, M Klein, M Wagner, et al. 2001. "Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm." *Nature* 411 (6835) (May 17): 298-302. doi:10.1038/35077067.

Duhme-Klair, Anne-K. 2009. "From Siderophores and Self-Assembly to Luminescent Sensors: The Binding of Molybdenum by Catecholamides." *European Journal of Inorganic Chemistry* 2009 (25) (September): 3689-3701. doi:10.1002/ejic.200900416.

Eddy, S R. 1998. "Profile hidden Markov models." *Bioinformatics (Oxford, England)* 14 (9): 755-763.

Edgar, R. C. 2004. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic Acids Research* 32 (March 8): 1792-1797. doi:10.1093/nar/gkh340.

Eisen, J A, S W Smith, and C M Cavanaugh. 1992. "Phylogenetic relationships of chemoautotrophic bacterial symbionts of Solemya velum say (Mollusca: Bivalvia) determined by 16S rRNA gene sequence analysis." *Journal of Bacteriology* 174 (10) (May): 3416-3421.

Escolar, L, J Pérez-Martín, and V de Lorenzo. 1999. "Opening the iron box: transcriptional metalloregulation by the Fur protein." *Journal of Bacteriology* 181 (20) (October): 6223-6229.

Faulkner, D. John. 2001. "Marine natural products." *Natural Product Reports* 18: 1-49. doi:10.1039/b006897g.

Fischbach, Michael A., and Christopher T. Walsh. 2006. "Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms." *Chemical Reviews* 106 (August): 3468-3496. doi:10.1021/cr0503097.

Fishman, Brian. 2010. Gene regulation in the shipworm symbiont, Teredinibacter turnerae strain T7901. M.S. Thesis, Oregon Health & Science University.

Francis, Christopher A, Kathryn J Roberts, J Michael Beman, Alyson E Santoro, and Brian B Oakley. 2005. "Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean." *Proceedings of the National Academy of Sciences of the United States of America* 102 (41) (October 11): 14683-14688. doi:10.1073/pnas.0506625102.

Frueh, Dominique P, Haribabu Arthanari, Alexander Koglin, David A Vosburg, Andrew E Bennett, Christopher T Walsh, and Gerhard Wagner. 2008. "Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase." *Nature* 454 (7206) (August 14): 903-906. doi:10.1038/nature07162.

Fuchs, B M, F O Glöckner, J Wulf, and R Amann. 2000. "Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes." *Applied and Environmental Microbiology* 66 (8) (August): 3603-3607.

Furrer, Jason L., Douglas N. Sanders, India G. Hook-Barnard, and Mark A. McIntosh. 2002. "Export of the siderophore enterobactin in Escherichia coli: involvement of a 43 kDa membrane exporter." *Molecular Microbiology* 44 (May 23): 1225-1234. doi:10.1046/j.1365-2958.2002.02885.x.

Gómez-Valero, Laura, Mario Soriano-Navarro, Vicente Pérez-Brocal, Abdelaziz Heddi, Andrés Moya, José Manuel García-Verdugo, and Amparo Latorre. 2004. "Coexistence of Wolbachia with Buchnera aphidicola and a secondary symbiont in the aphid Cinara cedri." *Journal of Bacteriology* 186 (19) (October): 6626-6633. doi:10.1128/JB.186.19.6626-6633.2004.

Gosalbes, M. J., A. Lamelas, A. Moya, and A. Latorre. 2008. "The Striking Case of Tryptophan Provision in the Cedar Aphid Cinara cedri." *Journal of Bacteriology* 190 (June 27): 6026-6029. doi:10.1128/JB.00525-08.

Graf, J., and E. G. Ruby. 2000. "Novel effects of a transposon insertion in the Vibrio fischeri glnD gene: defects in iron uptake and symbiotic persistence in addition to nitrogen utilization." *Molecular Microbiology* 37 (1) (July): 168-179. doi:10.1046/j.1365-2958.2000.01984.x.

Greenfield, L.J., and C.E. Lane. 1953. "Cellulose digestion in Teredo." *The Journal of Biological Chemistry* 204 (2) (October): 669-672.

Gros, Olivier, Liliane Frenkiel, and Marcel Moueza. 1998. "Gill filament differentiation and experimental colonization by symbiotic bacteria in aposymbiotic juveniles of Codkia orbicularis (Bivalvia: Lucinidae)." *Invertebrate Reproduction and Development* 34: 219-231.

Guan, L. L., K. Kanoh, and K. Kamino. 2001. "Effect of Exogenous Siderophores on Iron Uptake Activity of Marine Bacteria under Iron-Limited Conditions." *Applied and Environmental Microbiology* 67 (4) (April): 1710-1717. doi:10.1128/AEM.67.4.1710-1717.2001.

Hall, Thomas A. 1999. "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." *Nucleic Acids Symposium Series* 41: 95-98.

Hallam, Steven J, Konstantinos T Konstantinidis, Nik Putnam, Christa Schleper, Yoh-ichi Watanabe, Junichi Sugahara, Christina Preston, José de la Torre, Paul M Richardson, and Edward F DeLong. 2006a. "Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum." *Proceedings of the National Academy of Sciences of the United States of America* 103 (48) (November 28): 18296-18301. doi:10.1073/pnas.0608549103.

Hallam, Steven J, Tracy J Mincer, Christa Schleper, Christina M Preston, Katie Roberts, Paul M Richardson, and Edward F DeLong. 2006b. "Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota." *PLoS Biology* 4 (4) (April): e95. doi:10.1371/journal.pbio.0040095.

Hallé, F, and J M Meyer. 1992. "Iron release from ferrisiderophores. A multi-step mechanism involving a NADH/FMN oxidoreductase and a chemical reduction by FMNH2." *European Journal of Biochemistry / FEBS* 209 (2) (October 15): 621-627.

Hansen, A. K., and N. A. Moran. 2011. “Aphid genome expression reveals host-symbiont cooperation in the production of amino acids.” *Proceedings of the National Academy of Sciences* 108 (7) (January): 2849-2854. doi:10.1073/pnas.1013465108.

Hantke, K, G Nicholson, W Rabsch, and G Winkelmann. 2003. “Salmochelins, siderophores of Salmonella enterica and uropathogenic Escherichia coli strains, are recognized by the outer membrane receptor IroN.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (7) (April 1): 3677-3682. doi:10.1073/pnas.0737682100.

Harris, Wesley R., Carl J. Carrano, and Kenneth N. Raymond. 1979. “Coordination chemistry of microbial iron transport compounds. 16. Isolation, characterization, and formation constants of ferric aerobactin.” *Journal of the American Chemical Society* 101 (May): 2722-2727. doi:10.1021/ja00504a038.

Harris, Wesley R., Carl J. Carrano, Stephen R. Cooper, Stephen R. Sofen, Alex E. Avdeef, James V. McArdle, and Kenneth N. Raymond. 1979. “Coordination chemistry of microbial iron transport compounds. 19. Stability constants and electrochemical behavior of ferric enterobactin and model complexes.” *Journal of the American Chemical Society* 101 (September): 6097-6104. doi:10.1021/ja00514a037.

Haygood, Margo G., Pamela D. Holt, and Alison Butler. 1993. “Aerobactin Production by a Planktonic Marine Vibrio sp.” *Limnology and Oceanography* 38 (5) (July 1): 1091-1097. doi:10.2307/2838099.

Hewson, Ian, Rachel S Poretsky, Sonya T Dyhrman, Brian Zielinski, Angelicque E White, H James Tripp, Joseph P Montoya, and Jonathan P Zehr. 2009. “Microbial community gene expression within colonies of the diazotroph, Trichodesmium, from the Southwest Pacific Ocean.” *The ISME Journal* 3 (July 2): 1286-1300. doi:10.1038/ismej.2009.75.

Hibbing, Michael E, Clay Fuqua, Matthew R Parsek, and S Brook Peterson. 2010. “Bacterial competition: surviving and thriving in the microbial jungle.” *Nature Reviews. Microbiology* 8 (1) (January): 15-25. doi:10.1038/nrmicro2259.

Hoegy, F., X. Lee, S. Noel, D. Rognan, G. L. A. Mislin, C. Reimmann, and I. J. Schalk. 2009. “Stereospecificity of the Siderophore Pyochelin Outer Membrane Transporters in Fluorescent Pseudomonads.” *Journal of Biological Chemistry* 284 (22) (March): 14949-14957. doi:10.1074/jbc.M900606200.

Hoffmann, K, E Schneider-Scherzer, H Kleinkauf, and R Zocher. 1994. “Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis.” *Journal of Biological Chemistry* 269 (17) (April 29): 12710-12714.

Holmes, Bradley, and Harvey Blanch. 2006. “Genus-specific associations of marine sponges with group I crenarchaeotes.” *Marine Biology* 150 (July 4): 759-772. doi:10.1007/s00227-006-0361-x.

Hori, K, Y Yamamoto, T Minetoki, T Kurotsu, M Kanda, S Miura, K Okamura, J Furuyama, and Y Saito. 1989. “Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene.” *Journal of Biochemistry* 106 (4) (October): 639-645.

Hoyer, Katharina M, Christoph Mahlert, and Mohamed A Marahiel. 2007. “The iterative gramicidin s thioesterase catalyzes peptide ligation and cyclization.” *Chemistry & Biology* 14 (1) (January): 13-22. doi:10.1016/j.chembiol.2006.10.011.

Huber, T., G. Faulkner, and P. Hugenholtz. 2004. “Bellerophon: a program to detect chimeric sequences in multiple sequence alignments.” *Bioinformatics* 20 (April 8): 2317-2319. doi:10.1093/bioinformatics/bth226.

Ito, Yusai, and Alison Butler. 2005. “Structure of synechobactins, new siderophores of the marine cyanobacterium Synechococcus sp. PCC 7002.” *Limnology and Oceanography* 50: 1918-1923. doi:10.4319/lo.2005.50.6.1918.

Jones, David T., William R. Taylor, and Janet M. Thornton. 1992. “The rapid generation of mutation data matrices from protein sequences.” *Bioinformatics* 8: 275-282. doi:10.1093/bioinformatics/8.3.275.

Jukes, Thomas H., and Charles R. Cantor. 1969. *Evolution of protein molecules*. Mammalian Protein Metabolism. Academic Press, New York.

Keating, T A, D E Ehmann, R M Kohli, C G Marshall, J W Trauger, and C T Walsh. 2001. “Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis.” *Chembiochem: A European Journal of Chemical Biology* 2 (2) (February 2): 99-107.

Keating, T A, Z Suo, D E Ehmann, and C T Walsh. 2000. “Selectivity of the yersiniabactin synthetase adenylation domain in the two-step process of amino acid activation and transfer to a holo-carrier protein domain.” *Biochemistry* 39 (9) (March 7): 2297-2306.

Keller, Laurent, and Michael G. Surette. 2006. “Communication in bacteria: an ecological and evolutionary perspective.” *Nature Reviews Microbiology* 4 (February 27): 249-258. doi:10.1038/nrmicro1383.

Kester, Dana R., Iver W. Duedall, Donald N. Connors, and Ricardo M. Pytkowicz. 1967. “Preparation of artificial seawater.” *Limnology and Oceanography* 12: 176-179. doi:10.4319/lo.1967.12.1.0176.

Kesty, Nicole C, Kevin M Mason, Mary Reedy, Sara E Miller, and Meta J Kuehn. 2004. “Enterotoxigenic Escherichia coli vesicles target toxin delivery into mammalian cells.” *The EMBO Journal* 23 (November 18): 4538-4549. doi:10.1038/sj.emboj.7600471.

Kiers, E Toby, Robert A Rousseau, Stuart A West, and R Ford Denison. 2003. “Host sanctions and the legume-rhizobium mutualism.” *Nature* 425 (6953) (September 4): 78-81. doi:10.1038/nature01931.

Klappenbach, J A, J M Dunbar, and T M Schmidt. 2000. “rRNA operon copy number reflects ecological strategies of bacteria.” *Applied and Environmental Microbiology* 66 (4) (April): 1328-1333.

Kloepper, Joseph W., John Leong, Martin Teintze, and Milton N. Schroth. 1980. “Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria.” *Nature* 286 (5776) (August): 885-886. doi:10.1038/286885a0.

Kulp, Adam, and Meta J. Kuehn. 2010. “Biological Functions and Biogenesis of Secreted Bacterial Outer Membrane Vesicles.” *Annual Review of Microbiology* 64 (October 13): 163-184. doi:10.1146/annurev.micro.091208.073413.

Küpper, Frithjof C, Carl J Carrano, Jens-Uwe Kuhn, and Alison Butler. 2006. “Photoreactivity of iron(III)-aerobactin: photoproduct structure and iron(III) coordination.” *Inorganic Chemistry* 45 (15) (July 24): 6028-6033. doi:10.1021/ic0604967.

Kustka, Adam B., Sergio A. Sañudo-Wilhelmy, Edward J. Carpenter, Douglas Capone, James Burns, and William G. Sunda. 2003. “Iron requirements for dinitrogen- and ammonium-supported growth in cultures of Trichodesmium (IMS 101): Comparison with nitrogen fixation rates and iron:carbon ratios of field populations.” *Limnology and Oceanography* 48 (5): 1869-1884. doi:10.4319/lo.2003.48.5.1869.

Lambalot, R H, A M Gehring, R S Flugel, P Zuber, M LaCelle, M A Marahiel, R Reid, C Khosla, and C T Walsh. 1996. “A new enzyme superfamily - the phosphopantetheinyl transferases.” *Chemistry & Biology* 3 (11) (November): 923-936.

Lane, D.J. 1991. 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, 115-175. E. Stackebrandt and M. Goodfellow (ed.). John Wiley & Sons, Inc., New York, N.Y.

Larkin, M A, G Blackshields, N P Brown, R Chenna, P A McGettigan, H McWilliam, F Valentin, et al. 2007. “Clustal W and Clustal X version 2.0.” *Bioinformatics (Oxford, England)* 23 (21) (November 1): 2947-2948. doi:10.1093/bioinformatics/btm404.

Lechene, C. P., Y. Luyten, G. McMahon, and D. L. Distel. 2007. “Quantitative Imaging of Nitrogen Fixation by Individual Bacteria Within Animal Cells.” *Science* 317 (5844) (September): 1563-1566. doi:10.1126/science.1145557.

Lee, Eun-Young, Hong Kum Lee, Yoo Kyung Lee, Chung Ja Sim, and Jung-Hyun Lee. 2003. “Diversity of symbiotic archaeal communities in marine sponges from Korea.” *Biomolecular Engineering* 20 (4-6) (July): 299-304.

Leininger, S, T Urich, M Schloter, L Schwark, J Qi, G W Nicol, J I Prosser, S C Schuster, and C Schleper. 2006. “Archaea predominate among ammonia-oxidizing prokaryotes in soils.” *Nature* 442 (7104) (August 17): 806-809. doi:10.1038/nature04983.

Litwin, C M, T W Rayback, and J Skinner. 1996. “Role of catechol siderophore synthesis in Vibrio vulnificus virulence.” *Infection and Immunity* 64 (7) (July): 2834-2838.

Loomis, Lawrence D., and Kenneth N. Raymond. 1991. “Solution equilibria of enterobactin and metal-enterobactin complexes.” *Inorganic Chemistry* 30 (March): 906-911. doi:10.1021/ic00005a008.

Loper, J E, and M D Henkels. 1999. “Utilization of heterologous siderophores enhances levels of iron available to Pseudomonas putida in the rhizosphere.” *Applied and Environmental Microbiology* 65 (12) (December): 5357-5363.

de Lorenzo, V, A Bindereif, B H Paw, and J B Neilands. 1986. “Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in Escherichia coli K-12.” *Journal of Bacteriology* 165 (2) (February): 570-578.

de Lorenzo, V, S Wee, M Herrero, and J B Neilands. 1987. “Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor.” *Journal of Bacteriology* 169 (6) (June): 2624-2630.

Ludwig, Wolfgang, Oliver Strunk, Ralf Westram, Lothar Richter, Harald Meier, Yadhukumar, Arno Buchner, et al. 2004. “ARB: a software environment for sequence data.” *Nucleic Acids Research* 32 (4): 1363-1371. doi:10.1093/nar/gkh293.

Luyten, Y. A., J. R. Thompson, W. Morrill, M. F. Polz, and D. L. Distel. 2006. “Extensive Variation in Intracellular Symbiont Community Composition among Members of a Single Population of the Wood-Boring Bivalve Lyrodus pedicellatus (Bivalvia: Teredinidae).” *Applied and Environmental Microbiology* 72 (1) (January): 412-417. doi:10.1128/AEM.72.1.412-417.2006.

Marfey, Peter, and Martin Ottesen. 1984. “Determination ofd-amino acids. I. Hydrolysis of DNP-l-amino acid methyl esters with carboxypeptidase-Y.” *Carlsberg Research Communications* 49 (6) (January): 585-590. doi:10.1007/BF02908687.

Margot, H., C. Acebal, E. Toril, R. Amils, and J. Fernandez Puentes. 2002. “Consistent association of crenarchaeal Archaea with sponges of the genus Axinella.” *Marine Biology* 140 (April 1): 739-745. doi:10.1007/s00227-001-0740-2.

Mark Moore, C., Matthew M. Mills, Eric P. Achterberg, Richard J. Geider, Julie LaRoche, Mike I. Lucas, Elaine L. McDonagh, et al. 2009. “Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability.” *Nature Geoscience* 2 (November 1): 867-871. doi:10.1038/ngeo667.

Martín, J F. 1998. “New aspects of genes and enzymes for beta-lactam antibiotic biosynthesis.” *Applied Microbiology and Biotechnology* 50 (1) (July): 1-15.

Martin, Jessica D, Yusai Ito, Vanessa V Homann, Margo G Haygood, and Alison Butler. 2006. “Structure and membrane affinity of new amphiphilic siderophores produced by Ochrobactrum sp. SP18.” *Journal of Biological Inorganic Chemistry: JBIC: A Publication of the Society of Biological Inorganic Chemistry* 11 (5) (July): 633-641. doi:10.1007/s00775-006-0112-y.

Martinez, J S, G P Zhang, P D Holt, H T Jung, C J Carrano, M G Haygood, and A Butler. 2000. “Self-assembling amphiphilic siderophores from marine bacteria.” *Science (New York, N.Y.)* 287 (5456) (February 18): 1245-1247.

Martinez, Jennifer S, Jayme N Carter-Franklin, Elizabeth L Mann, Jessica D Martin, Margo G Haygood, and Alison Butler. 2003. “Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (7) (April 1): 3754-3759. doi:10.1073/pnas.0637444100.

Mashburn, Lauren M., and Marvin Whiteley. 2005. “Membrane vesicles traffic signals and facilitate group activities in a prokaryote.” *Nature* 437 (September 15): 422-425. doi:10.1038/nature03925.

May, J J, T M Wendrich, and M A Marahiel. 2001. “The dhb operon of Bacillus subtilis encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin.” *The Journal of Biological Chemistry* 276 (10) (March 9): 7209-7217. doi:10.1074/jbc.M009140200.

Miethke, M., and M. A. Marahiel. 2007. “Siderophore-Based Iron Acquisition and Pathogen Control.” *Microbiology and Molecular Biology Reviews* 71 (September 5): 413-451. doi:10.1128/MMBR.00012-07.

Miller, Mark A., Wayne Pfeiffer, and Terri Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop*, 1-8. New Orleans, LA.

Mills, Matthew M., Celine Ridame, Margaret Davey, Julie La Roche, and Richard J. Geider. 2004. “Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic.” *Nature* 429 (May 20): 292-294. doi:10.1038/nature02550.

Milton, D L, R O’Toole, P Horstedt, and H Wolf-Watz. 1996. “Flagellin A is essential for the virulence of Vibrio anguillarum.” *Journal of Bacteriology* 178 (5) (March): 1310-1319.

Mootz, Henning D, Dirk Schwarzer, and Mohamed A Marahiel. 2002. “Ways of assembling complex natural products on modular nonribosomal peptide synthetases.” *Chembiochem: A European Journal of Chemical Biology* 3 (6) (June 3): 490-504. doi:10.1002/1439-7633(20020603)3:6<490::AID-CBIC490>3.0.CO;2-N.

Moya, Andrés, Juli Peretó, Rosario Gil, and Amparo Latorre. 2008. “Learning how to live together: genomic insights into prokaryote–animal symbioses.” *Nature Reviews Genetics* 9 (March): 218-229. doi:10.1038/nrg2319.

Neilands, J B, T J Erickson, and W H Rastetter. 1981. “Stereospecificity of the ferric enterobactin receptor of Escherichia coli K-12.” *The Journal of Biological Chemistry* 256 (8) (April 25): 3831-3832.

Neufeld, D S, and S H Wright. 1995. “Basolateral transport of taurine in epithelial cells of isolated, perfused Mytilus californianus gills.” *The Journal of Experimental Biology* 198 (Pt 2) (February): 465-473.

O’Brien, I G, and F Gibson. 1970. “The structure of enterochelin and related 2,3-dihydroxy-N-benzoylserine conjugates from Escherichia coli.” *Biochimica Et Biophysica Acta* 215 (2) (August 14): 393-402.

Ohkuma, M. 2008. “Symbioses of flagellates and prokaryotes in the gut of lower termites.” *Trends in Microbiology* 16 (7) (July): 345-352. doi:10.1016/j.tim.2008.04.004.

Patzer, S. I., and V. Braun. 2009. “Gene Cluster Involved in the Biosynthesis of Griseobactin, a Catechol-Peptide Siderophore of Streptomyces sp. ATCC 700974.” *Journal of Bacteriology* 192 (2) (November): 426-435. doi:10.1128/JB.01250-09.

Percival Zhang, Y-H, Michael E Himmel, and Jonathan R Mielenz. 2006. "Outlook for cellulase improvement: screening and selection strategies." *Biotechnology Advances* 24 (5) (October): 452-481. doi:10.1016/j.biotechadv.2006.03.003.

Persmark, M, D Expert, and J B Neilands. 1989. "Isolation, characterization, and synthesis of chrysobactin, a compound with siderophore activity from Erwinia chrysanthemi." *Journal of Biological Chemistry* 264 (6) (February 25): 3187 - 3193.

Piel, J. 2006. "Bacterial symbionts: prospects for the sustainable production of invertebrate-derived pharmaceuticals." *Current Medicinal Chemistry* 13 (1): 39-50.

Pollack, J R, and J B Neilands. 1970. "Enterobactin, an iron transport compound from Salmonella typhimurium." *Biochemical and Biophysical Research Communications* 38 (5) (March 12): 989-992.

Preston, C.M., K.Y. Wu, T.F. Molinski, and E.F. DeLong. 1996. "A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov." *Proceedings of the National Academy of Sciences of the United States of America* 93 (13) (June 25): 6241-6246.

Pruesse, Elmar, Christian Quast, Katrin Knittel, Bernhard M Fuchs, Wolfgang Ludwig, Jörg Peplies, and Frank Oliver Glöckner. 2007. "SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB." *Nucleic Acids Research* 35 (21): 7188-7196. doi:10.1093/nar/gkm864.

Quadri, Luis E.N., Jason Sello, Thomas A. Keating, Paul H. Weinreb, and Christopher T. Walsh. 1998. "Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin." *Chemistry & Biology* 5 (November): 631-645. doi:10.1016/S1074-5521(98)90291-5.

Raquel G. Soengas, Cristina Anta, Alfonso Espada, Vanessa Paz, Isabel R. Ares, Miguel Balado, Jaime Rodriguez, Manuel L. Lemos, and Carlos Jimenez. 2006. "Structural characterization of vanchrobactin, a new catechol siderophore produced by the fish pathogen Vibrio anguillarum serotype O2." *Tetrahedron Letters* 47: 7113-7116.

Rasmussen, R. 2001. Quantification on the LightCycler. In *Rapid Cycle Real-time PCR, Methods and Applications*, 21-34. Meuer S., Wittwer C., and Nakagawara K. (ed.). Springer Press, Heidelberg.

Ratledge, Colin, and Lynn G Dover. 2000. "Iron metabolism in pathogenic bacteria." *Annual Review of Microbiology* 54 (1) (October): 881-941. doi:10.1146/annurev.micro.54.1.881.

Rausch, Christian, Ilka Hoof, Tilmann Weber, Wolfgang Wohlleben, and Daniel H Huson. 2007. "Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution." *BMC Evolutionary Biology* 7: 78. doi:10.1186/1471-2148-7-78.

Rausch, Christian, Tilmann Weber, Oliver Kohlbacher, Wolfgang Wohlleben, and Daniel H. Huson. 2005. "Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector

machines (TSVMs).” *Nucleic Acids Research* 33 (October 12): 5799-5808. doi:10.1093/nar/gki885.

Raymond, Kenneth N, Emily A Dertz, and Sanggoo S Kim. 2003. “Enterobactin: an archetype for microbial iron transport.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (7) (April 1): 3584-3588. doi:10.1073/pnas.0630018100.

Raymond, Kenneth N., Gertraud Muller, and Berthold F. Matzanke. 1984. “Complexation of iron by siderophores a review of their solution and structural chemistry and biological function.” *Topics in Current Chemistry* 123: 49-102.

Rippka, R., J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier. 1979. “Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria.” *Microbiology* 111 (March 1): 1-61. doi:10.1099/00221287-111-1-1.

Rose, Timothy M, Jorja G Henikoff, and Steven Henikoff. 2003. “CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design.” *Nucleic Acids Research* 31 (13) (July 1): 3763-3766.

Rotthauwe, J H, K P Witzel, and W Liesack. 1997. “The ammonia monooxygenase structural gene amoA as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations.” *Applied and Environmental Microbiology* 63 (12) (December): 4704-4712.

Rottig, M., M. H. Medema, K. Blin, T. Weber, C. Rausch, and O. Kohlbacher. 2011. “NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity.” *Nucleic Acids Research* 39 (May 9): W362-W367. doi:10.1093/nar/gkr323.

Rubio, Luis M., and Paul W. Ludden. 2008. “Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase.” *Annual Review of Microbiology* 62 (1) (October): 93-111. doi:10.1146/annurev.micro.62.081307.162737.

Rue, Eden L., and Kenneth W. Bruland. 1995. “Complexation of iron(III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method.” *Marine Chemistry* 50 (August): 117-138. doi:10.1016/0304-4203(95)00031-L.

Sakurai, Makiko, Ryuichi Koga, Tsutomu Tsuchida, Xian-Ying Meng, and Takema Fukatsu. 2005. “Rickettsia symbiont in the pea aphid Acyrthosiphon pisum: novel cellular tropism, effect on host fitness, and interaction with the essential symbiont Buchnera.” *Applied and Environmental Microbiology* 71 (7) (July): 4069-4075. doi:10.1128/AEM.71.7.4069-4075.2005.

Samel, Stefan A, Björn Wagner, Mohamed A Marahiel, and Lars-Oliver Essen. 2006. “The thioesterase domain of the fengycin biosynthesis cluster: a structural base for the macrocyclization of a non-ribosomal lipopeptide.” *Journal of Molecular Biology* 359 (4) (June 16): 876-889. doi:10.1016/j.jmb.2006.03.062.

Sandy, Moriah, and Alison Butler. 2009. “Microbial Iron Acquisition: Marine and Terrestrial Siderophores.” *Chemical Reviews* 109 (10) (October): 4580-4595. doi:10.1021/cr9002787.

Sandy, Moriahs, and Alison Butler. 2011. “Chrysobactin siderophores produced by Dickeya chrysanthemi EC16.” *Journal of Natural Products* 74 (5) (May 27): 1207-1212. doi:10.1021/np200126z.

Sandy, Moriah, Andrew Han, John Blunt, Murray Munro, Margo Haygood, and Alison Butler. 2010. “Vanchrobactin and Anguibactin Siderophores Produced by Vibrio sp. DS40M4.” *Journal of Natural Products* 73 (6) (June 25): 1038-1043. doi:10.1021/np900750g.

Scarrow, Robert C., David J. Ecker, Chiu. Ng, Sylvia. Liu, and Kenneth N. Raymond. 1991. “Iron(III) coordination chemistry of linear dihydroxyserine compounds derived from enterobactin.” *Inorganic Chemistry* 30 (March): 900-906. doi:10.1021/ic00005a007.

Schloss, P. D., and J. Handelsman. 2005. “Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness.” *Applied and Environmental Microbiology* 71 (March 3): 1501-1506. doi:10.1128/AEM.71.3.1501-1506.2005.

Sharp, K. H., B. Eam, D. J. Faulkner, and M. G. Haygood. 2007. “Vertical Transmission of Diverse Microbes in the Tropical Sponge Corticium sp.” *Applied and Environmental Microbiology* 73: 622-629. doi:10.1128/AEM.01493-06.

Shaw-Reid, C A, N L Kelleher, H C Losey, A M Gehring, C Berg, and C T Walsh. 1999. “Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of E. coli EntF catalyzes both elongation and cyclolactonization.” *Chemistry & Biology* 6 (6) (June): 385-400. doi:10.1016/S1074-5521(99)80050-7.

Shaw, Allison K, Aaron L Halpern, Karen Beeson, Bao Tran, J Craig Venter, and Jennifer B H Martiny. 2008. “It’s all relative: ranking the diversity of aquatic bacterial communities.” *Environmental Microbiology* 10 (9) (September): 2200-2210. doi:10.1111/j.1462-2920.2008.01626.x.

Shimodaira, H, and M Hasegawa. 2001. “CONSEL: for assessing the confidence of phylogenetic tree selection.” *Bioinformatics (Oxford, England)* 17 (12) (December): 1246-1247.

Shimodaira, Hidetoshi. 2002. “An approximately unbiased test of phylogenetic tree selection.” *Systematic Biology* 51 (3) (June): 492-508. doi:10.1080/10635150290069913.

Simon, N, V Coulanges, P Andre, and D J Vidon. 1995. “Utilization of exogenous siderophores and natural catechols by Listeria monocytogenes.” *Applied and Environmental Microbiology* 61 (4) (April): 1643-1645.

Simon, R., U. Priefer, and A. Pühler. 1983. “A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria.” *Bio/Technology* 1 (9) (November): 784-791. doi:10.1038/nbt1183-784.

Sipe, A R, A E Wilbur, and S C Cary. 2000. “Bacterial symbiont transmission in the wood-boring shipworm Bankia setacea (Bivalvia: Teredinidae).” *Applied and Environmental Microbiology* 66 (4) (April): 1685-1691.

Sneath, Peter H.A., and Robert R. Sokal. 1973. *Numerical taxonomy*. W.H. Freeman and Company, San Francisco.

Stachelhaus, Torsten, Henning D Mootz, and Mohamed A Marahiel. 1999. "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases." *Chemistry & Biology* 6 (8) (August): 493-505. doi:10.1016/S1074-5521(99)80082-9.

Stahl, D.A., and R. Amann. 1991. Development and application of nucleic acid probes. In *Nucleic acid techniques in bacterial systematics*, 205-248. E. Stackebrandt and M. Goodfellow (ed.). John Wiley & Sons, Ltd., Chichester, England.

Stamatakis, A. 2006. "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." *Bioinformatics* 22 (August 23): 2688-2690. doi:10.1093/bioinformatics/btl446.

Steger, Doris, Piers Ettinger-Epstein, Stephen Whalan, Ute Hentschel, Rocky de Nys, Michael Wagner, and Michael W Taylor. 2008. "Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges." *Environmental Microbiology* 10 (4) (April): 1087-1094. doi:10.1111/j.1462-2920.2007.01515.x.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods." *Molecular Biology and Evolution* 28 (May 4): 2731-2739. doi:10.1093/molbev/msr121.

Tavare, Simon. 1986. "Some probabilistic and statistical problems in the analysis of DNA sequences." *Lectures on Mathematics in the LIfe Sciences* 17 (57-86).

Taylor, Michael W, Regina Radax, Doris Steger, and Michael Wagner. 2007. "Sponge-associated microorganisms: evolution, ecology, and biotechnological potential." *Microbiology and Molecular Biology Reviews: MMBR* 71 (2) (June): 295-347. doi:10.1128/MMBR.00040-06.

Treusch, Alexander H, Sven Leininger, Arnulf Kletzin, Stephan C Schuster, Hans-Peter Klenk, and Christa Schleper. 2005. "Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling." *Environmental Microbiology* 7 (12) (December): 1985-1995. doi:10.1111/j.1462-2920.2005.00906.x.

Tseng, Claire C, Steven D Bruner, Rahul M Kohli, Mohamed A Marahiel, Christopher T Walsh, and Stephan A Sieber. 2002. "Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase." *Biochemistry* 41 (45) (November 12): 13350-13359.

Turner, Ruth D. 1966. *A Survey and Illustrated Catalogue of the Teredinidae*. The Museum of Comparative Zoology, Harvard University, Cambridge, MA.

Vraspir, Julia M, and Alison Butler. 2009. "Chemistry of marine ligands and siderophores." *Annual Review of Marine Science* 1: 43-63.

van Wageningen, AndréM.A., Peter N. Kirkpatrick, Dudley H. Williams, Barbara R. Harris, Jo K. Kershaw, Nicola J. Lennard, M. Jones, Steven J.M. Jones, and Patricia J. Solenberg. 1998. "Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic." *Chemistry & Biology* 5 (March): 155-162. doi:10.1016/S1074-5521(98)90060-6.

Wagner, Björn, Stephan A Sieber, Marcus Baumann, and Mohamed A Marahiel. 2006. "Solvent engineering substantially enhances the chemoenzymatic production of surfactin." *ChemBioChem* 7 (4) (April): 595-597. doi:10.1002/cbic.200500495.

Wallner, G, R Amann, and W Beisker. 1993. "Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms." *Cytometry* 14 (2): 136-143. doi:10.1002/cyto.990140205.

Walsh, Christopher T., Jun Liu, Frank Rusnak, and Masahiro Sakaitani. 1990. "Molecular studies on enzymes in chorismate metabolism and the enterobactin biosynthetic pathway." *Chemical Reviews* 90 (7) (November): 1105-1129. doi:10.1021/cr00105a003.

Waterbury, J. B., C. B. Calloway, and R. D. Turner. 1983. "A Cellulolytic Nitrogen-Fixing Bacterium Cultured from the Gland of Deshayes in Shipworms (Bivalvia: Teredinidae)." *Science* 221 (4618) (September): 1401-1403. doi:10.1126/science.221.4618.1401.

Watson, Robert J., Susan A. Joyce, Georgette V. Spencer, and David J. Clarke. 2005. "The exbD gene of Photorhabdus temperata is required for full virulence in insects and symbiosis with the nematode Heterorhabditis." *Molecular Microbiology* 56 (3) (March): 763-773. doi:10.1111/j.1365-2958.2005.04574.x.

Webster, N S, J E Watts, and R T Hill. 2001. "Detection and phylogenetic analysis of novel crenarchaeote and euryarchaeote 16S ribosomal RNA gene sequences from a Great Barrier Reef sponge." *Marine Biotechnology (New York, N.Y.)* 3 (6) (November): 600-608. doi:10.1007/s10126-001-0065-7.

Wichard, Thomas, Jean-Philippe Bellenger, Aurélie Loison, and Anne M. L. Kraepiel. 2008. "Catechol Siderophores Control Tungsten Uptake and Toxicity in the Nitrogen-Fixing Bacterium Azotobacter vinelandii." *Environmental Science & Technology* 42 (7) (April): 2408-2413. doi:10.1021/es702651f.

Winkelmann, G. 2002. "Microbial siderophore-mediated transport." *Biochemical Society Transactions* 30 (4) (August): 691-696. doi:10.1042/.

Wolf, M. K., and J. H. Crosa. 1986. "Evidence for the Role of a Siderophore in Promoting Vibrio anguillarum Infections." *Microbiology* 132 (10) (October): 2949-2952. doi:10.1099/00221287-132-10-2949.

Wu, Jingfeng, and George W. Luther. 1995. "Complexation of Fe(III) by natural organic ligands in the Northwest Atlantic Ocean by a competitive ligand equilibration method and a kinetic approach." *Marine Chemistry* 50 (August): 159-177. doi:10.1016/0304-4203(95)00033-N.

Wuchter, Cornelia, Ben Abbas, Marco J L Coolen, Lydie Herfort, Judith van Bleijswijk, Peer Timmers, Marc Strous, et al. 2006. "Archaeal nitrification in the ocean." *Proceedings of the National Academy of Sciences of the United States of America* 103 (33) (August 15): 12317-12322. doi:10.1073/pnas.0600756103.

Yang, Joyce C., Ramana Madupu, A. Scott Durkin, Nathan A. Ekborg, Chandra S. Pedamallu, Jessica B. Hostetler, Diana Radune, et al. 2009. "The Complete Genome of Teredinibacter turnerae T7901: An Intracellular Endosymbiont of Marine Wood-Boring Bivalves (Shipworms)." *PLoS ONE* 4 (7) (July): e6085. doi:10.1371/journal.pone.0006085.

Zhang, Guangping, Shady A. Amin, F pper, Pamela D. Holt, Carl J. Carrano, and Alison Butler. 2009. "Ferric Stability Constants of Representative Marine Siderophores: Marinobactins, Aquachelins, and Petrobactin." *Inorganic Chemistry* 48 (December 7): 11466-11473. doi:10.1021/ic901739m.