ABSTRACT

Title of Document:	ISOLATION AND CHARACTERIZATION OF A SPONGE-ASSOCIATED ACTINOMYCETE THAT PRODUCES MANZAMINES
	Olivier Peraud, Doctor of Philosophy, 2006
Directed By:	Professor Russell T. Hill Marine Estuarine Environmental Sciences

Two Indonesian sponges, *Acanthostrongylophora* sp. Sponge 35 and Sponge 52, containing manzamine A were collected off the coast of Manado, Indonesia. Manzamines are a family of marine alkaloids that exhibit a complex molecular architecture and possess bioactivities including antitumour, antimicrobial, antiparasitic and insecticidal activities. Manzamines have been found in 17 different species of sponges with wide geographical distribution which has led to speculation that they may be produced by a microbial symbiont rather than by the sponges themselves.

The sponges' microbial communities were investigated using 16S rRNA gene analysis and a rational culture-based microbiology approach in which specific bacterial groups were targeted. The molecular analysis of these microbial communities revealed that they were complex and diverse. Microbiological analyses

were conducted on *Acanthostrongylophora* sp. with a particular emphasis on the isolation of actinomycetes because of the high number of actinomycete sequences in this sponge 16S rRNA gene clone library and their excellent track record as bioactive compound producers.

One of the isolated actinomycetes, *Micromonospora* sp. strain M42, produces manzamine A and 8-hydroxy-manzamine, compounds initially detected in the sponge. A detailed analysis of *Micromonospora* sp. strain M42 showed that it grew on a wide range of salt concentrations with an optimal growth at 0-1% NaCl. Cultures of *Micromonospora* sp. strain M42 consistently produced manzamine A with a maximum yield of 1 mg/l. The genome size of *Micromonospora* sp. strain M42 was estimated at 6.7 Mb by pulsed field gel electrophoresis. The biosynthetic gene pathway encoding manzamine A was investigated using both biochemistry and molecular methods yet it remains elusive. *Micromonospora* sp. strain M42 underwent UV mutagenesis leading to isolation of mutants with yield of manzamine A improved by 3.5 fold. One of the mutants produces manzamine B, the putative biosynthetic precursor of manzamine A. A fosmid library of *Micromonospora* sp. strain M42 was constructed and low-pass genome sequencing gave insights into the strain's genome and revealed a high number of genes devoted to the production of secondary metabolites including polyketides and non-ribosomal peptides.

The isolation of *Micromonospora* sp. strain M42 greatly improves the chances of manzamines becoming a drug class for treatment of malaria.

ISOLATION AND CHARACTERIZATION OF A SPONGE-ASSOCIATED
ACTINOMYCETE THAT PRODUCES MANZAMINES

By

Olivier Peraud

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor Russell T. Hill, Ph.D., Chair
Professor Mark T. Haman, Ph.D.
Professor Frank T. Robb, Ph.D.
Professor Feng Chen, Ph.D.
Assistant Investigator Jacques Ravel, Ph.D.

Dedication

This dissertation is dedicated to my parents,

Jeannine and Michel

Statement of Contribution

Dr. K. V. Rao (University of Mississippi) performed the manzamine assays on samples from *Micromonospora* sp. strain M42 cultures.

Ms. Jennifer Allman (University of Mississippi) assisted in providing figures of the molecular structures.

Mr. Matthew Anderson (Center of Marine Biotechnology) mutated *Micromonospora* sp. strain M42 and screened the mutants for hyperproducing mutants.

Dr. Jacques Ravel (The Institute for Genomic Research) contributed to the construction of *Micromonospora* sp. strain M42 fosmid library. The sequencing of *Micromonospora* sp. strain M42 fosmid library was done at TIGR thanks to Dr. Ravel who also helped in the sequence analysis.

Acknowledgements

I would like to thank my advisor, Russell Hill who gave me the opportunity to work on this truly exciting research project. He has been a great help all along and taught me discipline and rigor in my work. Russell greatly improved my English both in speech and writing. I enjoyed our arguments and the fact that I was always treated as a colleague even when I was a junior graduate student. Finally, Russell, thank you for having been there when I had difficult times, especially at the beginning.

Next, I would like to thank my committee members, Mark, Frank, Feng and Jacques who have always been very accessible and contributed greatly to this work with their suggestions and analyses.

I would like to especially thank Jacques Ravel for his help and also for having listened to me complaining in French. Playing table tennis at TIGR was lots of fun and perhaps I will beat you someday.

Matthew Anderson has been very involved in this project and carried out the mutation studies of *Micromonospora* sp. strain M42. Thank you for your help and for putting up with all my always urgent last minute requests.

I would also like to thank Julie Enticknap. She's been a good friend and her stories were always a laugh.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BLAST   Basic Alignment Search Tool

bp   base pair

DGGE   Denaturing Gradient Gel Electrophoresis

DNA   deoxyribonucleic acid

EDTA   ethylenediaminetetraacetic acid

FISH   fluorescent *in situ* hybridisation

HPLC   High Performance Liquid Chromatography

kb   kilobase

LCMS   Liquid Chromatography Mass Spectrometry

LC-TOF-MS   Liquid Chromatography-Time Of Flight-Mass Spectrometry

Mb   megabase

NMR   Nuclear Magnetic Resonance

NRPS   non ribosomal peptide synthetase

PCR   Polymerase Chain Reaction

PFGE   Pulsed Field Gel Electrophoresis

PKS   polyketide synthase

SDS   sodium dodecyl sulfate

TAE   Tris acetate EDTA

TBE   Tris boric acid EDTA

1 Introduction

1.1 Sponges

Sponges have long been familiar organisms and were used commonly during antiquity by Romans and Greeks. They used *Spongia officinalis adriatica* essentially for bathing purposes but ancient Greek soldiers also used sponges as lining for their armor and helmets (Smith, 1875). There are records of sponges being used by Arabic physicians as early as 932 A.D. Sponges soaked with narcotic drugs were placed over the patient's nose to provide a state of anesthesia (Syed, 2002).

Sponges were considered plants until the mid 1700s when Peter Pallas and John Ellis challenged that popular opinion by proposing that sponges were in fact animals despite their sessility and absence of muscular contractions. A sponge phylum, Porifera, was eventually created, named after the porous surface of their tissue (Grant, 1836).

Sponges are one of the oldest metazoans dating back to Precambrian times. Sponge fossils have been found in Precambrian rock deposits in South China and dated at 580 million years (Li *et al.*, 1998). These very well conserved fossils contained soft tissues, amoebocytes and metazoan embryos. Cambrian rock deposits provided more than a thousand sponge fossils representing 15 genera and 30 species showing the early radiation of this phylum. Nowadays, the number of different sponge species is estimated at 15,000, found principally in tropical reefs but also at polar latitudes as well as in fresh water lakes and rivers (Berquist, 1978).

1.1.1 Sponge biology

A sponge can be defined as a sedentary, filter-feeder metazoan that has neither organs nor true tissues. Sponges are organized around a system of pores, ostia, canals and chambers that are used to canalize the large flow of water that is pumped through sponges. The water enters the sponge through the inhalant canals and exits by the oscules. It is thought that the ancestors of sponges included protozoans such as choanoflagellates, ameboid cells and siliceous- or calcareous-producing organisms that provided the skeletal material. Colonies of these protozoan ancestors would have exchanged their genetic material to eventually result in the first Precambrian sponge (Wilkinson, 1992).

A sponge is constituted of three layers. The first layer comprises pinacocytes and is called the pinacoderm. Under the pinacoderm is the mesohyl region that contains canals and choanocyte chambers. This is where the sponge metabolism, reproduction and nutrient transfer occur. The third layer is the choanoderm and contains choanocytes. The choanocyte cells line the choanocyte chambers. They are flagellated cells possessing a collar of cytoplasmic tentacles. It is through the movement of these tentacles that the flow of water is created bringing in nutrients. Sponges feed unselectively on particles present in the water up to 50 μm in size which is about the maximum size that the pores allow. Phagocytosis of the particles is done by archeocytes before the nutrients make their way to the mesohyl. The estimated number of choanocyte chambers is between 7000 and 18000 per cubic millimeter of sponge with each chamber pumping approximately 1200 times its own volume in water each day (Wilkinson, 1992).

The reproduction of sponges can be sexual or asexual. Some sponge species are hermaphrodite while others are permanently or temporarily male or female. Cross-fertilization occurs, though it has been rarely observed. Sperm are released into the exhalant stream, traveling with the water stream to the next sponge. Sperm are then trapped by choanocytes and transferred to the eggs inside the mesohyl. For viviparous species, eggs are released in the exhalant stream or in the water by dissolution of the dermal membrane (Berquist, 1978). Sponges can release larvae that are ready to settle and start the metamorphosis into young sponges. Larvae are ciliated and some can have swimming capabilities. Larvae have shown phototaxis and geotaxis raising the possibility that sponge distribution may not be only the result of selective mortality following the settlement. Settlement can be achieved within 20 hours for sponge species like *Halichondria moorei*. During the settlement phase, the larva rests on its anterior pole and the cells of this region spread evenly over the substrate. Within 24 hours, the larva loses its ciliated ring and the canal system with a single apical osculum becomes functional. The size of the young sponge at this stage is 600 μm and it will keep its single osculum and hemispherical shape until it reaches 2.0 mm diameter after 2 weeks.

1.1.2 Sponge taxonomy

The classification of sponges is quite difficult especially as it is primarily based on the physical observation of the animals (shape, color and type of spicules) rather than being resolved at a molecular level. Spicules are an important part of the sponge skeleton and because of their diversity in size and shape, constitute recognizable characters for the

identification of sponges. The name of spicules varies depending on the number of axes and the number of rays. In references to axes, spicules can be mono, di-, tri- or tetraxonic. When rays are refered to, a spicule can be di-, tri-, tetra, penta or hexactinic (Bergquist, 1978). About 7000 species have been described and it is estimated that this is only half the number of species still to be discovered. The phylum Porifera is composed of three classes, the Calcarea (calcareous sponges), the Hexactinellida (glass sponges) and the Desmospongiae (keratose sponges).

The calcareous sponges have a mineral skeleton composed entirely of calcium carbonate, with free, rarely linked or cemented , di-, tri-, tetra- and/or polyactinal spicules, occasionally with a solid basal calcitic skeleton (Manuel *et al.*, 2002). The number of calcareous sponge species described represents less than 5% of all described sponges. The Calcarea class is composed of 2 subclasses, 5 orders, 22 families and 75 genera (Manuel *et al.*, 2002). Most species are quite unremarkable being usually small and colorless. Many calcareous sponges live in cryptic habitats like marine caves, overhangs or interstices of hard substrata. Calcareous sponges are viviparous and their larvae are blastulae (an early embryonic form produced by cleavage of a fertilized ovum and consisting of a spherical layer of cells surrounding a fluid-filled cavity).

Hexactinellida, also known as glass sponges, are deep-water marine sponges. They are defined by the production of siliceous spicules of hexactinic, triaxonic symmetry, or shapes derived from this symmetry. They have neither calcareous minerals nor sclerified spongin as skeletal components. Their spicules can be entirely loose or partially fused to

form a rigid basal and choanosal framework. There are about 500 described species of Hexactinellids. The Hexactinellida class is composed of 5 orders, 17 family and 118 genera (Reiswig, 2002). These sponges are viviparous and the only species that was studied produced distinctive trichimella larvae. Recently, glass sponges have attracted much interest for both their structural complexity (Aizenberg *et al.*, 2005) and the ability of their spicules to conduct light more efficiently than man-made optic fibers (Sundar *et al.*, 2003). The study of the structural complexity of *Euplectella* sp. showed a textbook example in mechanical engineering with seven hierarchical organization levels in the sponge skeleton and construction strategies such as laminated structures, fiber-reinforced composites, bundled beams and diagonally reinforced square-grid cells (Aizenberg *et al.*, 2005).

Demospongiae account for about 85% of all living sponges representing about 6000 described species with estimates of at least twice that number of species still to be discovered. Demospongiae contain discrete cellular elements and have parenchymella (stereoblastula) or blastula larvae. They can be either viviparous or oviparous. The siliceous spicules composing their skeleton are monaxonic or tetraxonic bound together with spongin in discrete fibers or they can be loosely aggregated with collagenous filaments forming the ground substance of the intracellular matrix (Hooper & Van Soest, 2002). They never have triaxonic spicules. Demosponges come in all sort of shapes. They can be encrusted, massive, lobate, tubular, branching, flabellate or cup-shaped. There are a few freshwater species of demosponges but this class is essentially marine with habitats ranging from the intertidal to the deepest seas. The demospongiae class is composed of 15 orders, 88 families and 500 valid genera (Hooper & Van Soest, 2002).

1.1.3 Sponge microbiology

Sponges have developed a complex association with a very diverse range of microbes including bacteria, cyanobacteria, dinoflagelates, diatoms and archaea. Although the proportion of sponge-associated bacteria can vary dramatically from sponge to sponge, bacteria can occupy up to 60% of the sponge volume, being densely packed in the intercellular and intracellular matrix (Wilkinson, 1978a). In the sponge *Lamellodysidea herbacea*, *Oscillatoria spongelliae*, a cyanobacterial symbiont, accounts for 50% of the sponge cellular volume (Berthold *et al.*, 1982) and in a *Xestospongia* sp. sponge from the Indian Ocean, eubacteria represented as much as 56% of the sponge biomass (Brantley *et al.*, 1995). The large number of bacteria present in the sponges often exceeds the amount of bacteria present in the water column by one to three orders of magnitude (Webster & Hill, 2001).

Microbes associated with sponges can be divided into three categories: exosymbiont, endosymbiont and intracellular symbiont. The exosymbionts are present on the outer layers of the sponge while the endosymbionts are found in the mesohyl. The intracellular symbionts are present within the sponge cells and sometime inside the nuclei of sponge cells (Fuerst *et al.*, 1999; Lee *et al.*, 2001; Vacelet, 1970). A distinction needs to be made between sponge-associated bacteria and bacteria as food particles. Because of the intense filtering of the surrounding water, sponges accumulate a number of bacteria traveling with the stream of water. These bacteria are present in the choanocyte chambers and are not part of the sponge microbial community. How sponges acquire symbiotic bacteria is still unclear in most cases but two hypotheses have been proposed. The first hypothesis

is that symbionts come from the filtered water. They would have to resist the digestion and immune response from the sponge and migrate out of the canal system to the mesohyl. Electron micrographs have shown that some sponge-associated bacteria have thickened cell walls or slime capsules that may help them to resist digestion by sponge archeocytes (Friedrich *et al.*, 1999). The second hypothesis is the vertical transmission of symbionts through larvae, eggs or buds. This has been demonstrated in the case of the ubiquitous alpha-proteobacteria sponge symbiont group related to NW001 (Webster & Hill, 2001). Organisms closely related to NW001 were found in seven geographically distant sponge genera and the study of larvae from *Mycale laxissima* showed that a NW001-like strain was present in the larvae, providing evidence for a vertical transmission of the symbiont (Enticknap *et al.*, 2006).

Bacterial symbionts are believed to provide their host sponge with a range of benefits: nutrient acquisition (Borowitzka *et al.*, 1988), stabilization of the sponge skeleton (Wilkinson *et al.*, 1981), processing of metabolic waste (Wilkinson, 1978c), protection from UV light (Shick & Dunlap, 2002) and chemical defense (Schmidt *et al.*, 2000).

Early studies on the diversity of bacteria associated with sponges depended on the cultivation of the microbes as well as the microscopical observations of sponge sections. The culture-based studies from environmental samples are known for their limitations because of "the great plate count anomaly" (Staley & Konopka, 1985). The colonies that are easily grown on plates usually represent less than 1% of all microbial cells present in the sample. Though for sponges there are some rare examples, like in the case of the sponge *Ceratoporella nicholsoni*, where 3 to 11% of the total bacteria were culturable

(Santavy *et al.*, 1990), cultured sponge-associated bacteria generally follow the norm by representing less than 1% of the total sponge microbial cells being recovered as culturable colonies (Webster & Hill, 2001). This last decade saw molecular techniques being applied to the exploration and understanding of sponge-associated microbial communities. The use of community analysis 16S rRNA gene sequencing has overcome limitations associated with culture-based community studies and provided a detailed analysis of the sponge microbial community, revealing the presence of bacterial species never suspected before. The first use of this culture independent technique on the marine sponge *Discodermia* spp. revealed the presence of novel gamma-proteobacteria and other uncultivated strains (Lopez *et al.*, 1999). Another pioneer study using a similar approach was done on the Great Barrier Reef sponge *Rhopaloeides odorabile* and showed the remarkable microbial diversity in this sponge. Members of several bacterial classes constitute the microbial community of *R. odorabile* among which beta-proteobacteria, gamma-proteobacteria, *Cytophaga/Flavobacterium*, *Actinobacteria* as well as euryarcheotes and crenarcheotes (Webster *et al.*, 2001a; Webster *et al.*, 2001b). A comparative study of the microbial communities of the geographically distant sponges *Aplysina aerophoba*, *Theonella swinhoei* and *R. odorabile* suggested that there is a uniform microbial community among sponges from different oceans (Hentschel *et al.*, 2002). It is important to emphasize that sponge-associated microbial communities have been shown to be very diverse and very different from those of the surrounding water column.

1.2   Marine natural products

Marine natural products research is a multidisciplinary field where chemists, pharmacologists and biologists work together to extract, test and understand the bountiful molecules found in the oceans. The field of natural products is not recent but remains a strong developing field because of the size of the "playground" to explore as the oceans cover two thirds of the Earth. Marine natural products discovery includes the search for pharmaceuticals, enzymes, dietary supplements and biopolymers. The search for new pharmaceuticals is a US$20 billion effort per annum although only a small portion of these funds have been directed towards marine natural products discovery. The marine environment is attractive as it holds promise as a source of entirely new bioactive compounds. By isolating compounds from the marine environment, scientists may avoid the rediscovery problem that has been plaguing terrestrial natural product discovery.

Several thousand bioactive compounds have been isolated from marine macro- and microorganisms (Blunt *et al.*, 2003; Blunt *et al.*, 2004; Blunt *et al.*, 2005; Faulkner, 1986; Faulkner, 1987; Faulkner, 1988; Faulkner, 1990; Faulkner, 1991; Faulkner, 1993; Faulkner, 1994; Faulkner, 1996; Faulkner, 1998; Faulkner, 2001; Faulkner, 2002) and their activity tested. Marine natural products activities are very diverse and include anti-viral, anti-cancer, anti-parasitic, anti-HIV, anti-bacterial, anti-fungal, analgesic, anti-inflammatory activities. Many drugs from the sea are now in clinical trial such as bryostatin-1, a compound extracted from the bryozoan *Bugula neritina* in phase II clinical trials for treatment of various type of cancer like chronic lymphocyte and metastatic melanoma. Didemnin B is a cyclic depsipeptide compound isolated from the tunicate *Trididemnum solidum*. It is a potent anti-cancer compound also in phase II

9

clinical trials. Ecteinascidin-743 is a promising anti-tumor agent isolated from the

Caribbean mangrove tunicate *Ecteinascidin turbinata*. It acts by binding to the small

groove of DNA and alkylating the DNA, it is in phase II clinical trials. In 2004, 39

marine-derived natural products were in clinical or preclinical trials (Newman & Cragg,

2004). The first drug from the sea to have made it on the market is Prialt (ziconotide)

(Olivera *et al.*, 1985). It is a peptide isolated from the cone snail *Conus magnus* and is

used as a pain-killer. It is a thousand times more powerful than morphine without being

addictive. Prialt has been recently approved by the Food and Drug Administration on

December 28th 2004 and is commercialized by Elan Corp.

Marine sponges are undoubtedly the largest contributing phylum of natural products from

the sea. These sessile, filter feeding, animals have provided more than a thousand

different metabolites in the last five years alone (Blunt *et al.*, 2003; Blunt *et al.*, 2004;

Blunt *et al.*, 2005; Faulkner, 2001; Faulkner, 2002). The bioactive compounds are very

diverse in both structure and bioactivity. Among the type of sponge bioactive

compounds are polyketides, alkaloids, sterols, cyclic peptides and terpenes to cite just a

few. The bioactivities of these compounds are as diverse as their structures and include

anti-cancer, anti-parasitic, antiviral, anti-inflammatory and antibiotic activities as well as

herbicidal and antifouling properties (Osinga *et al.*, 1998). Some sponge orders have

attracted more interest than others because of their excellent track record as sources of

bioactive compounds e.g. the lithistid sponges. One of the lithistid sponges that has been

well studied is *Discodermia dissoluta* that contains discodermolide, a polyketide with an

antitumor activity greater than that of the important anticancer drug Pacitaxel (taxol)

(Haygood *et al.*, 1999).

1.2.1 The "supply problem"

Marine natural products are diverse and often difficult and expensive to synthesize. The amount of metabolite found in the source organism is rarely enough to get through clinical trials. Increasing the amount of compound by a massive harvest of the source organism is rarely a viable option because of the disastrous ecological impact. In the case of bryostatin, 13,000 kg of the bryozoan *B. neritina* had to be harvested in order to obtain only 18 g of bryostatin for clinical trials.

There is a true supply issue that needs to be overcome in order to meet the requirements of the demand for those compounds that become successful drugs. In some cases, it may be possible to chemically synthesize the compound but most of the time, the complexity of the molecules or the costs involved preclude this approach. Aquaculture is another option but it can be unreliable, and there are reports of diseases wiping out the entire production. In the event of the compound being produced by a microbe, isolation and culture of the bacteria can provide with a reliable source for the bioactive compound of interest and open a wide range of possible production improvements.

1.2.2 Natural products from sponge-associated bacteria

1.2.2.1 Actinomycetes

The Gram-positive bacteria are well known for having provided more bioactive compounds than any other group of bacteria, accounting for approximately two-thirds of

the world's naturally occurring antibiotics by the 1980s (Berdy, 1989; Okami & Hotta, 1988). Actinomycetes were generally considered soil bacteria until about a decade ago when it became clear that they were abundant and active in marine sediments (Jensen *et al.*, 1991; Moran *et al.*, 1995). The first marine actinomycete taxon, *Salinispora*, was eventually described in 2005 (Maldonado *et al.*, 2005).

The screening of actinomycetes for bioactive compounds continues although often plagued by the rediscovery of previously identified metabolites, slowing the isolation of new molecules. A model study of the genus *Streptomyces* estimated the total number of antimicrobial compounds that this genus is capable of producing to be of the order of 150,000 (Watve *et al.*, 2001), a tiny fraction of which have already been characterized.

Since the application of molecular approaches to study the biodiversity of sponge microbial communities, actinomycetes have been shown to be a major component of these communities (Hentschel *et al.*, 2002; Webster *et al.*, 2001b). This exciting discovery suggests that sponges may be an important source of novel actinomycetes for drug discovery. A little-studied sub-class of the *Actinobacteria*, the *Acidimicrobidae* was recently found in marine sponges (Montalvo *et al.*, 2005).

The *Actinobacteria* described below represent only a small sub-set of the remarkable diversity of *Actinobacteria* present in marine sponges and efforts to culture additional *Actinobacteria* from marine sponges are certainly warranted considering the outstanding track-record of this group as a source of drugs and their diversity within sponges.

Streptomyces sp. strain KM86-9B was isolated from an unidentified Korean sponge (Lee *et al.*, 1998). The crude extract displayed significant inhibition on topoisomerase I, a class of enzymes that changes the degree of DNA supercoiling by causing single-strand

breaks and re-ligation. Further analysis of the crude extract revealed six iso- and anteiso-fatty acids. Four of them, 14-methylpentadecenoic acid, iso-16:1, 12-methyltetradecanoic acid, anteiso-15:0, cyclopropane fatty acid and 14-methylhexadecanoic acid, anteiso-17:0 exhibited activity comparable to that of the topoisomerase inhibitor camptothecin.

Streptomyces sp. strain NI80 was obtained from an unidentified Japanese sponge (Imamura *et al.*, 1993). Analysis of the supernatant resulted in the identification of two new antimycins called urauchimycin A (**1**) and B (**2**), the first antimycins to possess a branched side-chain moiety. Urauchimycin A and B are active against the fungus *Candida albicans* at the concentration of 10 μg/ml, by inhibiting its morphological differentiation.
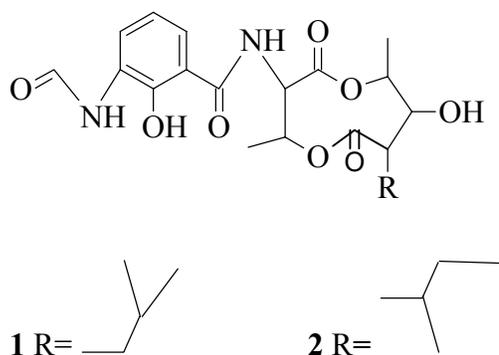
A *Saccharopolyspora* sp. nov. strain was isolated from the sponge *Mycale plumose* collected along the coast of Qingdao, China (Liu *et al.*, 2005). This strain demonstrated a potent *in vitro* anti-cancer activity against five cancer cell lines, P388, HL60, A-549, BEL-7402 and SPCA4. Analysis of the bioactive compounds produced by this *Saccharopolyspora* led to the isolation of metacycloprodigiosin and undecylprodigiosin. These prodigiosin analogs were already known but it is the first report from a

Saccharopolyspora sp. Prodigiosins are a family of polypyrrole red pigments that exhibits a wide range of biological activities such as immunosuppressive activities (Furstner *et al.*, 2001; Tsuji *et al.*, 1990), proton inhibition (Matsuya *et al.*, 2000; Sato *et al.*, 1998) and anti-cancer activity (Ermolyeva *et al.*, 1964; Montaner *et al.*, 2000). They are produced by several groups of microorganisms including *Streptomyces* and *Serratia*.

Micromonospora sp. strain L31-CLCO-002 was isolated from the homogenate of the sponge *Clathrina coriacea*, collected of the coast of Fuerteventure island in the Canary islands archipelago (Hernandez *et al.*, 2000). The culture of *Micromonospora* sp. strain L31-CLCO-002 yielded the isolation of two new indolocarbazole alkaloids, 4'-N-methyl-5'-hydroxystaurosporine and 5'-hydroxystaurosporine as well as staurosporine. Staurosporine posseses an inhibitory activity against fungi and yeast and inhibits protein kinase and platelet aggregation (Meksuriyen & Cordell, 1988).

Micromonospora sp. strain L-25-ES25-008 was isolated from an unidentified sponge collected off the coast of Mozambique in the Indian Ocean (Chimeno *et al.*, 2000). This strain produces a spiroketal 26-membered macrocyclic lactone structure related to oligomycins that showed cytotoxic activity against P-388, A-549, HT-29 and MEL-28, cell lines.

The Great Barrier sponge *Pseudoceratina clavata* yielded the isolation of ten bacterial strains (Kim *et al.*, 2005) related to the recently identified *Salinospora* group (Mincer *et al.*, 2002). The results of the bioassay conducted with these isolates showed activity against various bacteria, including *S. aureus* and *Staphylococcus epidermidis*. In a subquent study, Kim *et al.*(2006) conducted a phylogenetic analysis of the ketosynthase (KS) sequences of the previously isolated sponge associated *Salinispora* strains. The

results showed that the polyketide synthase sequences from the *Salinispora* strains were closely related to that of *Amycolatopsis mediterranei*, producer of rifamycin B. A chemical analysis of the *Salinispora* strains revealed that one of them, *Salinispora* M403 produces both rifamycin B **(5)** and SV **(6)**.

5   **6**
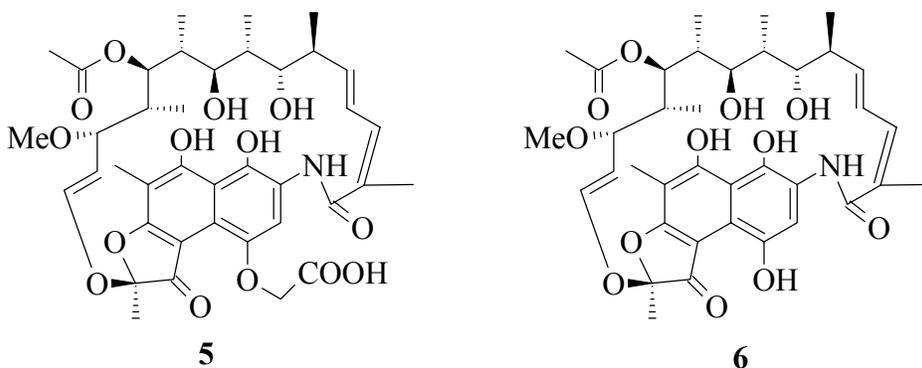
From the Mediterranean sponge *Halichondria panicea*, *Microbacterium* sp. strain HP2 was isolated (Wicke *et al.*, 2000). This actinomycete produces four cell associated glycoglycerolipids and one diphosphatidylglycerol. The main compound was identified as 1-*O*-acyl-3-[R-glucopyranosyl-(1-3)-(6-*O*-acyl-R-mannopyranosyl)]glycerol (GGL.2) and exhibited strong antitumor activity by inhibiting growth of the tumor cell lines HM02 and Hep G2 (50% inhibition at 0.4 to 3 μg/ml). By optimizing the culture condition and scaling up the production of *Microbacterium* sp. strain HP2, the yield of GGL 2 was improved to 300 mg/L (Lang *et al.*, 2004).

A *Micrococcus luteus* strain was isolated from a sponge *Xestospongia* sp. collected off Nouméa in New Caledonia (Bultel-Poncé *et al.*, 1997). The bacterium synthesized a novel acyl-1-(acyl-6'-mannobiosyl)-3-glycerol. The bioactivity of this glycoside was not tested. In a subsequent study, Bultel-Poncé *et al.*(1998) identified a 2,4,4'-trichloro-2'-hydroxyphenylether as well as acyl-1-(acyl-6'-mannobiosyl)-3-

glycerol from *M. luteus* isolated from *Xestospongia* sp. The chlorinated compound, that

was previously discovered in grapefruit seeds (Nishina *et al.*, 1991), showed

antimicrobial properties when tested against *S. aureus*, *Escherichia coli* and *Vibrio*

anguillarum.

An unidentified bacterium of the genus *Micrococcus* was isolated from the fire sponge

Tedania ignis (Stierle *et al.*, 1988). This isolate yielded four benzothiazoles, previously

undetected in *Tedania* spp, as well as three diketopiperazines already known from

terrestrial sources (Chen, 1960; Kodaira, 1961).

1.2.2.2 *Pseudomonas/Alteromonas*

The genus *Pseudomonas* comprises about 60 different species. They are Gram

negative bacteria, strict aerobes and mobile by polar flagella. *Pseudomonas* bacteria are

ubiquitous, they are found mainly in fresh and salt water but also in soil and plants. They

are known to produce a number of siderophores i.e. quinolobactin (Matthijs *et al.*, 2004)

and pyoverdin (Cox & Adams, 1985), antifungi i.e. ecomycins (Miller *et al.*, 1998) and

exotoxins (Heckly, 1964; Lynn & Callahan, 1974). *Alteromonas* are an essential

component of the marine environment. They can be found in various habitat including

coastal and open water areas (Ivanova & Mikhailov, 2001), deep-sea (Delong *et al.*,

1997), hydrothermal vents (Jeanthon, 2000) and sediments (Ivanova *et al.*, 2003).

A strain of *Pseudomonas aeruginosa* was isolated from the Antarctic sponge

Isodictya setifera (Jayatilake *et al.*, 1996). The bacterium was shown to produce six

diketopiperazines and two yellow phenazine alkaloids. Phenazine-1-carboxylic acid (**7**)

and phenazine-1-carboxamide (**8**) demonstrated antibacterial activity against *Bacillus*

cereus, *M. luteus* and *S. aureus*.
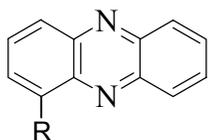
7 R= COOH   **8** R= CONH$_2$

An unidentified *Pseudomonas* or *Alteromonas* bacterium strain DF-1 was isolated from

the most common sponge of the Black Sea, *Dysidea fragilis* (De Rosa *et al.*, 2000). The

Black Sea is an especially interesting environment due to its lower salinity (18 ppt).

These peculiar characteristics led to the isolation of unusual compounds from algae and

invertebrates. Strain DF-1 was cultivated in Bacto-broth with a 16% salt concentration

matching that of the Black Sea. The analysis of the strain DF-1 culture revealed a novel

tripeptide (**9**) containing a β-aminopimelic acid.

9
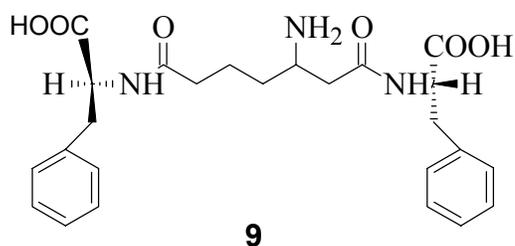
Pseudomonas sp. strain 1531-E7 was isolated from an *Homophymia* sp. sponge from

New Caledonia (Bultel-Poncé *et al.*, 1999). This gram negative bacterium produced four

quinolones and a phosphatidyl glyceride. Quinolone 1 (**10**) showed activities against

Plasmodium falciparum (ID$_{50}$ 1 µg/ml) and HIV-1 (ID$_{50}$ 10^{-3} µg/ml). Quinolone 2 (**11**)

demonstrated mild cytotoxicity against human epidermal carcinoma KB cells. Finally,

quinolone 4 (**12**) has antimicrobial activity against *S. aureus* and cytotoxicity toward KB cells (ID_{50} <2 µg/ml).
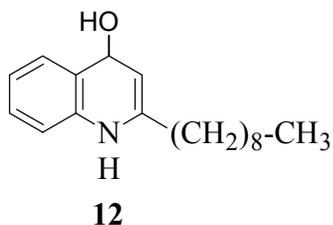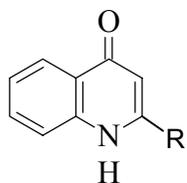
10 R= $(CH_2)_{10}$-CH_3
11 R= $(CH_2)_8$-CH_3

From the New Caledonia sponge *Suberea creba*, two bacteria identified as pseudomonad strains E14 and E15 were isolated (Debitus *et al.*, 1998). Extracts from these organisms' cultures were chromatographically identical and demonstrated low general toxicity and high antifouling properties. Further investigation leaded to the identification of four quinolones, phenazine-α-carboxamide, 2-n-heptylquinol-4-one, 3-n-heptyl-3-hydroxyquinolin-2,4-dione, a N-oxide-2-n-heptylquinolin derivative and a benzyldiketopiperazine.

Pseudomonas sp. strain KK10206C was cultivated from a homogenate of the sponge *Halichondria okadai* (Miki *et al.*, 1994). The compound okadaxanthine (**13**), a rare C_{50} carotenoid, was isolated from *Pseudomonas* sp. strain KK10206C. Carotenoids exhibit multiple biological properties, from light absorption acting as accessory pigments in photosynthetic organisms to antioxidants protecting against free radicals. Reported activities include animal reproductive improvement, decrease in cataract risk and anticancer properties (Krinsky, 1994).
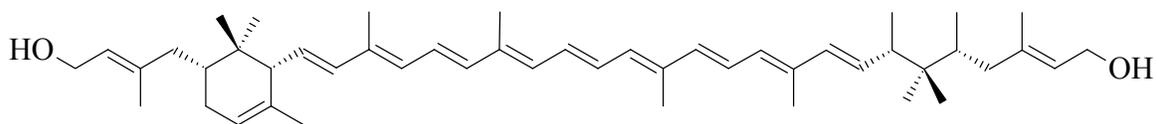
13

Pseudomonas sp. strain KP20-4 was isolated from the marine sponge *Cinachyrella australiensis*, collected in the Republic of Palau (Kanoh *et al.*, 2003) and found to produce two new siderophores, pseudoalterobactin A and B. Siderophores have a high affinity for iron (Fe^{3+}) making them iron scavengers. Iron is implicated in many biological reactions and is an essential element for most microorganisms. The concentration of dissolved iron in the ocean is low therefore the ability of a bacterium to produce siderophores may give the microbe a selective advantage in the marine environment. Siderophores have been evaluated as possible carriers for drug delivery system (Roosenberg *et al.*, 2000) and their immunosuppressive activity has been reported (Iijima *et al.*, 1999a; Iijima *et al.*, 1999b; Iijima *et al.*, 1999c).

The bacterium *Alteromonas* sp. isolated from the sponge *Halichondria okadai* collected off Nagai in Japan, produces the tetracyclic alkaloid alteramide A (**14**)(Shigemori *et al.*, 1992). Alteramide A is a macrocyclic lactam that exhibited cytotoxicity against murine leukemia P388 cells, murine lymphoma L1210 cells and the human epidermoid carcinoma KB cells.
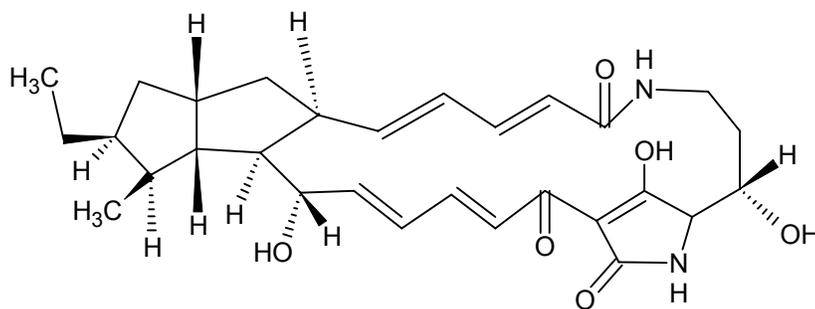
14

An A*lteromonas* bacterium also isolated from a *H. okadai* sponge, was found to produce

a ubiquinone (**15**) (Kon-ya *et al.*, 1995). This metabolite inhibits the settling of barnacle

larvae. Biofouling of submerged surfaces is a major concern in the marine environment,

for man-made structures such as pipeline, harbor piers and of course ships. There is an

urgent need to find substitutes for tri-butyl-tin (TBT) paints which have been banned due

to their toxic effect on the environment. For this reason, many efforts are focused on

marine invertebrates and including sponges that remain unfouled as potential sources of

antifouling compounds.

32

33 R=β–OH
34 R=H
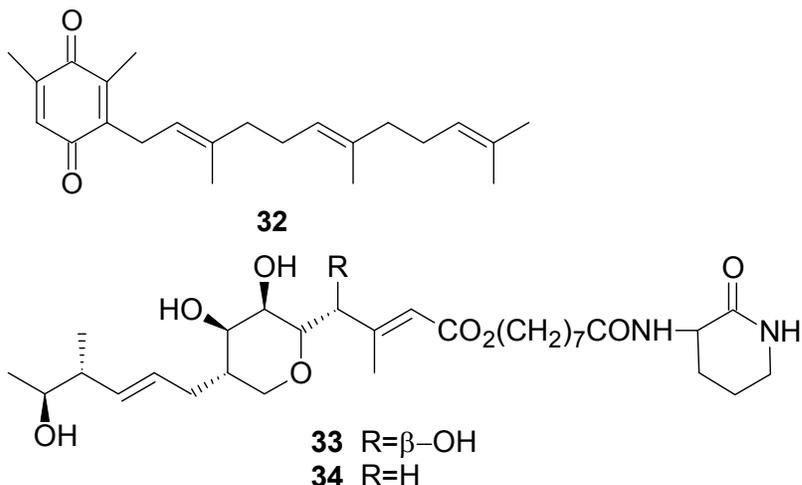
From the sponge *Darwinella rosacea*, collected in Harrington Sound, Bermuda, an

Alteromonas sp. was cultured on marine agar (Difco) (Stierle & Stierle, 1992). This pink

colored facultative anaerobe showed antimicrobial activity against *S. aureus,* later traced

back to two pseudomonic acids (**16,17**). Some pseudomonic acids such as pseudomonic

acid A, are currently used as drugs because of their antibacterial properties.

1.2.2.3 *Pseudoalteromonas*

Pseudoalteromonas are Gram negative bacteria found exclusively in the marine environment. They are often pigmented but non-luminescent. They are chemoorganotrophs, strict aerobes generally mobile by flagella (Ivanova *et al.*, 2002a). *Pseudoalteromonas* have a record of bioactive compounds i.e. active against antibiotic-resistant *S. aureus* (Isnansetyo & Kamei, 2003) and antifouling compounds that prevent the settling of organisms on surfaces such as pipelines or boats (Egan *et al.*, 2001; Venkateswaran & Dohmoto, 2000).

Pseudoalteromonas maricaloris strain KMM636[T] was isolated from the Australian sponge *Fascaplysinopsis reticulate* (Ivanova *et al.*, 2002b). Two pigments suspected to be novel monobrominated chromopeptides were isolated and showed antibacterial activity against *S. aureus*, *Enterococcus faecium*, *B. subtilis* and *Candida albicans*. The pigments have also cytotoxic activities against Ehrlich ascites-tumor cells and sea urchins eggs.

1.2.2.4 *Flexibacter*

Flexibacter are Gram negative bacteria that grow in a rod-shaped cell format. They are not flagellated but are motile by mean of gliding. *Flexibacter* is a known fish pathogen (Hansen *et al.*, 1992) which causes gill disease, as well as egg and larvae disease.

Flexibacter strains DK30213 and DK30223 were isolated the marine sponge *Reniera japonica*, collected from the Sagami Bay in Japan (Miki *et al.*, 1996). Both

strains were found to produce the carotenoid zeaxanthin. Carotenoids are known for their

free radical scavenging and quenching role. Miki *et al.* determined the total amount of

carotenoid in the sponge to be 7.8 mg/100 g of wet sponge. Six carotenoids were

detected among which zeaxanthin represented a substantial component. The authors

proposed that the host sponge would benefit from the bacterial carotenoid by using it as a

defense against free radicals generated by exposure to intense sunlight.

1.2.2.5 *Roseobacter/Ruegeria*

 Roseobacter are members of the alpha-proteobacteria class and representatives of

this clade are among the most readily cultured marine bacteria (Buchan *et al.*, 2005). A

notable member of this clade is the uncultured SAR83, the second most abundant

picoplancton after SAR11 (Beja *et al.*, 2002; Rappe *et al.*, 2000). Because of their

numbers and their ability to derive energy from light, the *Roseobacter* genus is believed

to be an important contributor of the estimated 5 to 10% energy generated in the upper

layer of tropical ocean (Goericke, 2002; Kolber *et al.*, 2001). Marine *Roseobacter* can be

found in diverse marine habitats such as sands, sediments, macroalgae (Shiba, 1991),

meromictic lakes (Yurkova *et al.*, 2002) but also associated with dinoflagellates (Lafay *et

al.*, 1995; Prokic *et al.*, 1998), in hot springs (Petursdottir & Kristjansson, 1997) and

sponges (Taylor *et al.*, 2004).

From the sponge *Cymbastela concentrica* was isolated an alpha-proteobacterium from the

Roseobacter-Ruegeria subgroup producing acyl homoserine lactone (AHL) (Taylor *et al.*,

2004). AHLs are used as signaling compounds by many gram-negative bacteria to

communicate with one another. These molecules can be responsible for modulation of

growth or attachment of bacteria. AHLs molecules are commonly found in biofilm and it

is no surprise to see such molecules in sponges where bacteria can be densely packed as

in biofilms.

A *Ruegeria* strain SDC-1 was isolated from cell cultures of *Suberites domuncula* (Mitova

et al., 2004b). This alpha proteobacterium was found to produce 10 diketopiperazines,

two of which, cyclo-(L-tyrosyl-*trans*-4-hydroxy-L-proline) and cyclo-(L-phenylalanyl-

trans-4-hydroxy-L-proline) were isolated for the first time. The authors hypothesized

that these ten cyclodipeptides were signal molecules of bacterial intraspecies and

interspecies quorum sensing. Mitova *et al.* (2004a) isolated eight different cyclic

peptides from the same *Ruegeria* strain SDC-1, with two of them being new cyclic

peptides. These eight cyclic peptides were different from the previously isolated ones.

All isolated cyclopeptides were tested for antimicrobial and antifungal activity. The two

new cyclopeptides, cyclo-(glycyl-L-seryl-L-glutamyl) (**18**) and cyclo-(glycyl-L-prolyl-L-

glutamyl) (**19**) showed moderate activity against *B. subtilis* (MIC of 25 and 50 µg/ml

respectively) while no activity was detected for the other cyclopeptides. The authors

speculated that these cyclopeptides could play some defensive role, protecting the

Ruegeria strain.

18

19
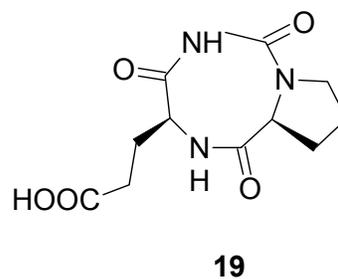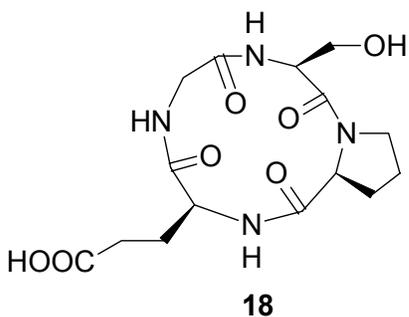
Cyclopeptide 3 was shown to form complexes with Ca^{2+} and Ba^{2+} cations but no function was hypothesized. Compounds 5 and 7 are known to be active as plant growth regulators while cyclopeptide 8 was shown to activate antibiotic production of *Pseudoalteromonas luteoviolacea* (Jiang *et al.*, 2000). Cyclopeptides 5-8 resemble cyclodipeptides that are able to interact with LuxR-mediated quorum sensing system of bacteria, making them potential actors in the cell to cell signaling system (Mitova *et al.*, 2004a).

1.2.2.6 Cyanobacteria

Cyanobacteria are well known for their abundance in the oceans and their role in both carbon and nitrogen cycles. Cyanobacteria are commonly found as dominant components associated with sponge microbial communities. Wilkinson (1983a) showed that in sponges containing cyanobacterial symbionts, primary production of the phototrophs was higher than the total respiration the sponge symbiont community. The sponge will largely benefit from this primary production, mostly through excreted metabolites like glycogen and glycerol (Wilkinson, 1983a). Cyanobacteria are capable of producing various bioactive secondary metabolites such as non-ribosomal peptides, polyketides and hybrid products of the two classes such as cryptophycin (**20**) and anti-inflammatory pigments like scytonemin (**21**) (Proteau *et al.*, 1993; Stevenson *et al.*, 2002).

20

21
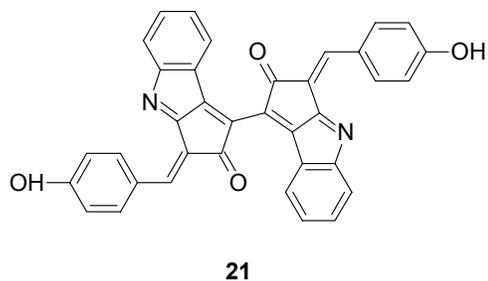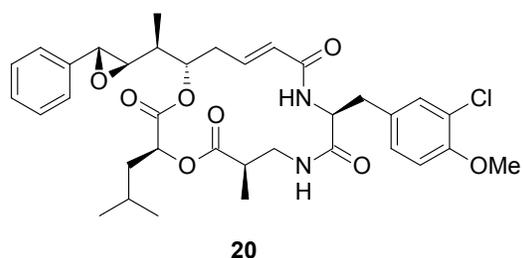
It has also been proposed that cyanobacteria may protect sponges from the UV radiation by producing mycosporine-like amino acids (Shick & Dunlap, 2002).

The filamentous cyanobacterium *Oscillatoria spongeliae* is found associated with the sponge *Lamellodysidea herbacea* (=*Dysidea herbacea*) and other dictyoracid sponges where it can constitute up to 50% of the sponge volume (Ridley *et al.*, 2005a). Using flow-cytometry on an Australian specimen of *L. herbacea*, Unson & Faulkner (1993), obtained a highly enriched cyanobacterial fraction (>95%) and attributed to *O. spongeliae* the production of 13-demethylisodysidenin. In a later study, using a similar approach on *L. herbacea* collected in Palau, Unson and colleagues identified the polybrominated biphenyl ether 2-(2',4'-dibromophenyl) 4,6-dibromophenol in the cyanobacterial fraction (Unson *et al.*, 1994). The brominated phenol showed activity against representative Gram negative and positive bacteria. This brominated compound deters fish feeding at or below natural concentrations found in the sponge (Paul, 1992). Flowers *et al.* (1998) investigated an Australian *L. herbacea*, using a Percoll density gradient to separate sponge cells from *O. spongeliae*, and showed that the diketopiperazine didechlorodihydrodysamide C and one analog were located within the cyanobacterium. A study by Ridley *et al.* (2005a) on four dictyoceratid sponges from Palau showed that each sponge species contains its own *O. spongeliae*, genetically different from one

25

another. Only one strain produced the chlorinated secondary metabolites dysidenin (**22**) and dysideathiazole.
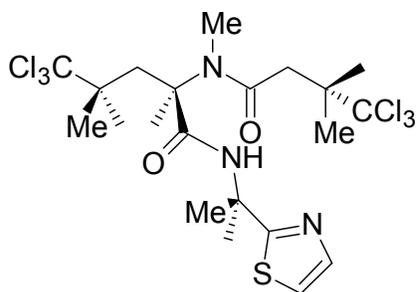
22

Within this strain, biosynthetic variability was observed indicated by the presence of one of two sets of biosynthetic genes, resulting in the strain being able to synthesize one or the other but not both chlorinated compounds. This observation strengthens previously formulated hypothesis about different strains of *O. spongeliae* being responsible for the production of the chlorinated and brominated compounds (Faulkner *et al.*, 1994). It is important to note that the symbiont *O. spongeliae* has not yet been cultured.
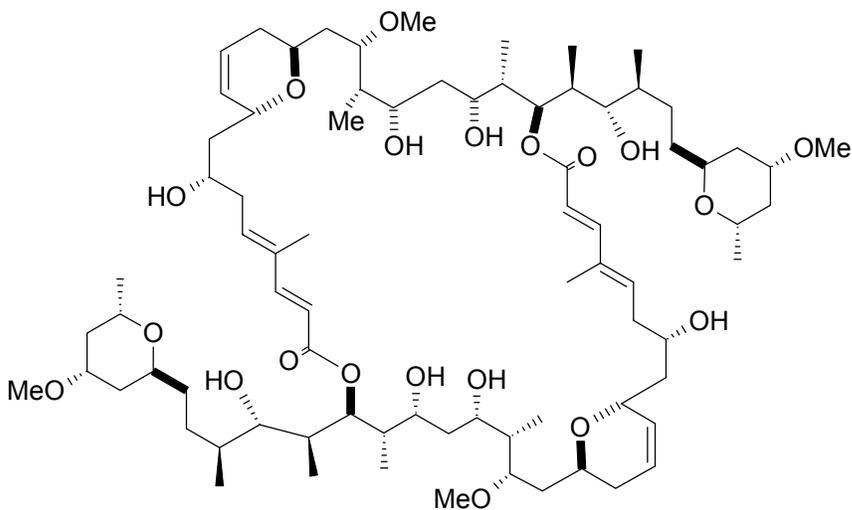
23

Recent work reports the isolation of swinholides from two different cyanobacteria (Andrianasolo *et al.*, 2005). Swinholide A (**23**) is a macrocyclic polyketide cytotoxin that has potent anti-cancer activity that results in the disruption of the actin cytoskeleton (Bubb *et al.*, 1995). It was originally isolated from the marine sponge *Theonella swinhoei* (Kitagawa *et al.*, 1990) but also found in other sponges of the *Theonella*, *Lamellomorpha* and *Tedenia* genera. The co-localization studies of swinholide A by Bewley *et al.* (1996) showed that the compound was associated with heterotrophic unicellular bacteria that were not cyanobacteria suggesting that the compound was not of cyanobacterial origin. The fact that cyanobacteria can indeed synthesize swinholide A, raises the possibility that the true metabolic origin of swinholides in sponges may be cyanobacterial although it is possible that the biosynthesis capability is shared by several different bacterial groups.

1.2.2.7 *Vibrio*

Two *Vibrio* sp. were isolated from an unidentified specimen of *Dysidea* sp., collected near the islands Tutuila and Ofu (Eastern Samoa) (Elyakov *et al.*, 1991). These bacteria contained brominated compound in the culture broth. Further investigation of *Vibrio* sp. strain KMM 9-81-1 resulted in the identification of 3,5 dibromo-2-(3',5'-dibromo-2'methoxyphenoxy) phenol. Since this report, extensive work has been done on *Dysidea* sp. symbiont *O. spongeliae* proving that it is also the source of brominated biphenyl ether (Faulkner *et al.*, 1994; Ridley *et al.*, 2005a; Unson & Faulkner, 1993). Though it is possible that *Vibrio* sp. strain KMM 9-81-1 shares with *O. spongeliae* the

ability to synthesize brominated biphenyl ether, Piel (2004) suggested that *Vibrio* sp.

strain KMM 9-81-1 may be a contaminant rather than a true symbiont of *Dysidea* sp.

From the sponge *Hyatella* sp., an unidentified *Vibrio* sp. was isolated and shown to

produce andrimid, an antibacterial mixed polyketide-peptide that was interestingly also

detected in the sponge (Oclarit *et al.*, 1994).

1.2.2.8 Sponge metabolites of likely bacterial origin

In addition to swinholide A, a second metabolite from the Paluan sponge *T.*

swinhoei, theopalauamide (**24**), an antifungal bicyclic glycopeptide, has been studied

(Schmidt *et al.*, 1998). Four distinct cell populations were consistently associated with

the sponge and were separated by centrifugation. Theopalaumide occurred in a fraction

composed of filamentous heterotrophic bacteria (Bewley *et al.*, 1996).
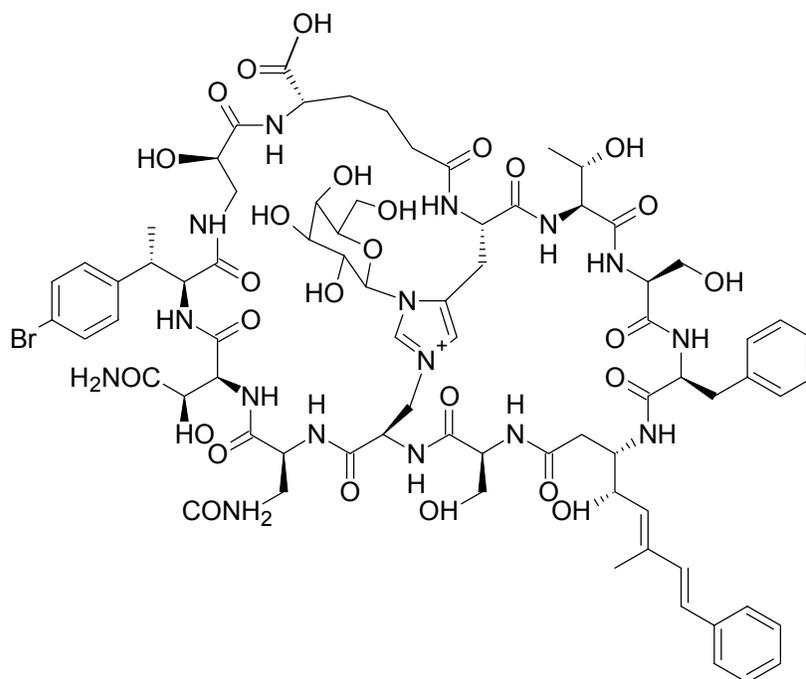
24

Further investigation by Schmidt *et al.* (2000) determined the 16S rRNA gene sequence of the filamentous symbiont using denaturing gradient gel electrophoresis, PCR and *in situ* hybridization. This symbiont belongs to the δ-proteobacteria class and because it was not cultured, was proposed as "*Candidatus* Entotheonella palauensis".

Mimosamycin (**25**) was isolated from the Fijian sponge *Xestospongia caycedoi* (McKee & Ireland, 1987). Mimosamycin is a neutral antibiotic mainly active against *Mycobacterium*. Interestingly, this compound is known to be produced by *Streptomyces lavendulae* n° 314 (Arai *et al.*, 1976). It is therefore quite legitimate to presume that origin of mimosamycin in the Fijian sponge is bacterial and the bacterium probably is an actinomycete.
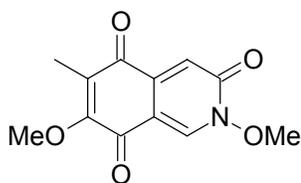
25

1.2.2.9 Future approaches

In some cases, it may be possible to directly access the genes responsible for the

biosynthesis of a product of interest without first culturing the "producer" symbiont. The

first example of the successful use of such approach is work by Piel *et al.* (2004) on a

polyketide from the pederin family, onnamide A (**26**) found in the sponge *T. swinhoei*.
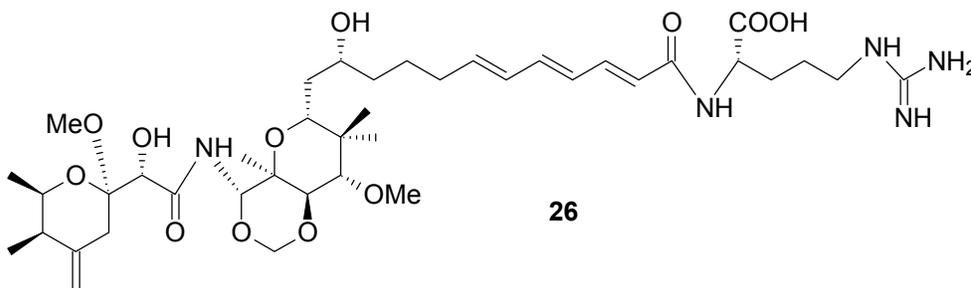
26

Building on a previous metagenomic approach to identify the bacterial origin of the

highly antitumor polyketide pederin found in rove beetles of the genera *Paederus* and

Paederidus (Piel, 2002), the authors constructed a metagenomic library of total DNA

from the sponge *T. swinhoei* and its associated microbes. This metagenomic library was

screened using polyketide synthase PCR primers and the genes responsible for the

synthesis of onnamide A were successfully isolated. Through the analysis of these genes,

it was concluded that the onnamide A genes were most likely of bacterial origin but the

precise identification of the bacterial source of the genes remains to be identified.

Using a metagenomic approach, Schirmer *et al.* (2005) generated a metagenomic library

of more than four gigabases for the sponge *D. dissoluta*. The authors were especially

interested in this sponge because of the potent antitumor compound discodermolide (**27**)

that was isolated from it (Gunasekera *et al.*, 1990). Discodermolide, because of its

molecular structure, is believed to be a polyketide synthesized by a bacterial type I

modular PKS.

27
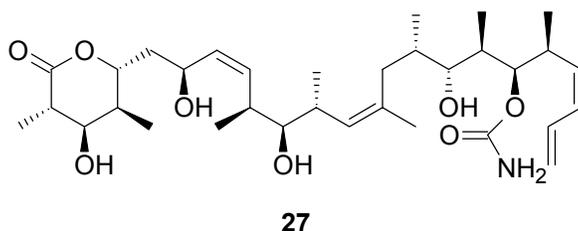
The authors' rationale was to screen the metagenomic library by PCR using degenerate

bacterial polyketide synthase primers in order to identify the PKS's gene cluster present

in the sponge-associated microbial community. The identification of a PKS cluster

consistent with the biosynthesis of discodermolide remained unsuccessful, nonetheless,

Schirmer *et al.* were able to assess the diversity of PKS and NRPS genes within *D.

dissoluta* microbial community. From the screening of the metagenomic library, 85

different KS sequences were reported and 17 KS domains were found to be most

homologous to cyanobacterial or myxobacterial type I PKS genes. Of the clones

analyzed, 50-68% had PKSs smaller than 35 kb, or 5 modules assuming an average

module size of 5 kb. Large multimodular PKSs were also found and the authors

described a PKS with a size of 110 kb coding for 25,572 amino acids with a predicted

molecular mass of ~2.7 MDa. This PKS comprises one starter module and fourteen

complete extender modules and one incomplete extender module. Schirmer *et al.* (2005)

predicted that this PKS would be a multimethyl-branched C_{30} fatty acid rather than a complex polyketide.

These reports demonstrate the power of the metagenomic approach in revealing genes of sponge-associated bioactive compounds. However, it may be difficult to match the genes with the bacterial host unless a 16S rRNA gene or other signature sequences are present on the cloned insert. It is interesting to note that these metagenomic libraries may be biased towards bacterial DNA as bacterially-derived DNA appeared to account for about 90% of the total sequences in the library. Another explanation is that this dominance of bacterial sequences is the result of the high numbers and great diversity of bacteria within the sponge.

Undoubtedly, metagenomic represents an important addition in the research arsenal for discovering biosynthesis pathways of sponge-associated bioactive metabolite. With only two reports of the application of metagenomic to the identification of sponge metabolites, the results are so far mixed, with success in isolating the biosynthesis pathway for onamide A (Piel *et al.*, 2004) and failure in isolating the discodermolide biosynthesis pathway (Schirmer *et al.*, 2005). As metagenomic becomes more widely applied in the field of marine natural products discovery, this approach may facilitate the discovery of hundreds of new metabolites, possibly identifying entirely new classes of compounds.

1.3 The actinomycetes

Actinomycetes are Gram-positive bacteria with DNA rich in guanine and cytosine. They are unicellular filamentous microorganisms that branch monopodially, more rarely

dichotomously. These filaments can be either of a single type called substrate or vegetative, or of two types, substrate and aerial. Some actinomycetes, like *Mycobacterium*, do not form mycelia and grow as pleomorphic or cocoid elements.

Ferdinand Cohn isolated and described, from the tear duct of an eye, the first actinomycetes in 1875 and in 1878, C.O. Hartz found a parasitic microbe in the infected jaw of a cow that he named *Actinomyces bovis*. Due to their filamentous aspect, actinomycetes were thought to be fungi, explaining the origin of the name actinomycetes which in Greek means "radiant fungi". Actinomycetes used to form a group on their own between the bacteria and the fungi but in the 1950s, after investigation of their chemical composition and fine structure, they were confirmed as prokaryotes and joined the bacterial domain.

Actinomycetes belong to the class *Actinobacteria* (Stackebrandt *et al.*, 1997), order *Actinomycetales* (Buchanan, 1917) which includes 10 suborders and 30 families. This relatively recent *Actinobacteria* class was proposed based on the 16S rDNA analysis of hundreds of actinomycete sequences.

1.3.1 Ecology of actinomycetes

Actinomycetes are found in a wide range of habitats. They are present in the frozen soils of Polar regions and in the dry soils of deserts. They can be found in crude oil, heavily metal-contaminated soil and sediments and fresh and salt water environments. They are not extremophiles and seem to be absent in highly acidic (pH<1) and extremely hot (hot spring) environments. Actinomycetes are mostly saprophytes

though some can form parasitic or symbiotic associations with animals and plants. Selman Waksman in the early part of the 20th century contributed greatly to the understanding of actinomycete ecology by publishing more than two hundred papers and many books on the subject and established the predominance of actinomycetes in soil (Waksman & Curtis, 1916; Waksman & Curtis, 1918). The techniques described in these studies along with those from Stanley Williams (Williams *et al.*, 1983; Williams *et al.*, 1984), another major contributor to the field of actinomycete ecology, are still in use in today's laboratories. In the last quarter of the 20th century, investigation of the marine environment such as near-shore and deep-sea sediments, have revealed the presence of actinomycetes (Jensen *et al.*, 1991; Weyland, 1969; Weyland, 1981). It is worth mentioning that despite the fact that oceans cover 70% of the Earth surface and contain among the most diverse ecosystems on the planet, they have not been widely recognized as an important source for novel actinomycetes. The distributions of actinomycetes in the marine environment and their ecological roles remain largely undescribed. For a long time, the existence of indigenous populations of marine actinomycetes was challenged. Actinomycetes produce resistant spores that can remain viable but dormant for many years and it was argued that the actinomycetes recovered from the marine environment were in fact the result of spores from soil actinomycetes that had washed into the oceans. This theory persisted despite evidence that actinomycetes can be recovered from deep-sea sediments (Weyland, 1969) and that marine actinomycetes can be metabolically active (Moran *et al.*, 1995) and physiologically adapted to the salt concentration encountered in the sea (Jensen *et al.*, 1991; Mincer *et al.*, 2002). *Rhodococcus marinonascens* was the first actinomycete species that was described and accepted as an autochthonous marine

species. Mincer *et al.* (2002) studied 212 actinomycete isolates from a group called Mar 1. These bacteria were isolated from geographically distant sediments collected from tropical or subtropical locations. The strains were distinguished by morphological characteristics, small-subunit rRNA gene signature nucleotides and by an obligate requirement for sea water for growth. Phylogenetic analysis of 16S rRNA gene sequences of seven strains showed that they formed a monophyletic clade within the family *Micromonosporaceae* suggesting novelty at the genus level. The Mar 1 strains were provisionally called '*Salinospora*'. Later the taxon was formally named *Salinispora*, belonging to the family *Micromonosporaceae* (Maldonado *et al.*, 2005).

1.3.2 Sponge-associated actinomycetes

Both sponges and actinomycetes are known for their excellent track record regarding the production of bioactive metabolites. With bacteria constituting up to 60% of the wet weight of some sponges (Wilkinson, 1978a), much effort has been devoted into culturing actinomycetes from sponges. Actinomycetes are resilient bacteria often found among culturable sponge microbes and members of genera such as *Streptomyces* (Imamura *et al.*, 1993; Lee *et al.*, 1998), *Micromonospora* (Peraud *et al.*, 2002), *Saccharopolyspora* (Liu *et al.*, 2005), *Gordonia* (Montalvo *et al.*, 2005), *Micrococcus* (Montalvo *et al.*, 2005), *Bradybacterium* (Montalvo *et al.*, 2005) and *Salinispora* (Kim *et al.*, 2005) have been isolated from sponges. With the broadening use of 16S rRNA gene sequence analyses for the study of sponge-associated microbial communities, actinomycetes have appeared as a consistent, diverse and sometime major component of

sponge-associated microbial communities (Hentschel *et al.*, 2002; Webster *et al.*, 2001b).

A study analyzing the microbial communities of *Xestospongia muta* from the Florida

Keys and *Xestospongia testudinaria* from Indonesia, showed the diversity of

actinomycetes within both sponges and highlighted Acidimicrobidae, a subclass of

Actinobacteria as a major component of the sponge's associated microbial communities

(Montalvo *et al.*, 2005). The different 16S rRNA gene sequence analyses of sponge

associated microbial communities have demonstrated that sponges are a good source for

actinomycetes both in abundance and diversity. These sponge-derived actinomycetes

could possibly be a source for new bioactive compounds. The role of sponge-associated

actinomycetes remains unclear. It is possible that by their saprophytic nature they are

involved in the processing of metabolic waste. The ability of actinomycetes to produce

secondary metabolites is also possibly helpful in protecting the host sponge against

predators, diseases and fouling.

1.3.3 Secondary metabolites

Actinomycetes are the most prolific microorganisms for the production of

antibiotics accounting for approximately two-thirds of the world's naturally occurring

antibiotics by the 1980s (Berdy, 1989; Okami & Hotta, 1988). It was in Selman A.

Waksman's laboratory that the first actinomycetes antibiotics were isolated, actinomycin,

an anti-tumor compound isolated in 1940 (Waksman & Woodruff, 1940), streptomycin a

potent antimicrobial compound in 1944, neomycin in 1949 and the anti-fungal compound

candicidin in 1953. The genus *Streptomyces* has been the main contributor of antibiotics

and two streptomycetes, *Streptomyces avermitilis* (Ikeda *et al.*, 2003) and *Streptomyces coelicolor* (Bentley *et al.*, 2002) have had their genomes sequenced. *S. avermitilis* is known to produce the insecticide avermectin that affects the nervous system of insects. The genome size of *S. avermitilis* is 9,025,608 bp and was found to contain 32 secondary metabolite clusters encoding pigments, siderophores, terpenes, peptides and polyketides representing a total size of 561,317 bp (16% of the total genome size) (Omura *et al.*, 2001). *S. coelicolor* has a genome size of 8,667,507 bp and contains more than 20 secondary metabolite clusters (Bentley *et al.*, 2002). Other actinomycete genera such as *Actinoplanes*, *Actinomadura* and *Micromonospora* produce pharmaceutically interesting secondary metabolites (Williams, 1989). Recently, rediscovery of existing secondary metabolites has been a problem (Zaehner & Fiedler, 1995) that slowed the isolation of new bioactive compounds. Actinomycetes, because of their excellent track record regarding secondary metabolites, remain one of the favored microbes in drug discovery. The sequencing of the genomes of *S. avermitilis* and *S. coelicolor* confirmed the unsurpassed potential of streptomycetes and actinomycetes in general, to produce useful secondary metabolites.

1.3.4 The genus *Micromonospora*

1.3.4.1 Description of the genus *Micromonospora*

The genus *Micromonospora* is a member of the family *Micromonosporaceae* belonging to the suborder *Micromonosporineae*, in the order *Actinomycetales* (Stackebrandt *et al.*, 1997). *Micromonospora* species are Gram-positive bacteria that are

sensitive to pH below 6.0 and grow between 20˚ and 40˚C. *Micromonospora* have a septate mycelium averaging 0.5 μm in diameter and non-motile spores borne singly, sessile on short or long sporophores that often occur in branched clusters. There is no aerial mycelium. The walls of *Micromonospora* contain meso-diaminopimelic acid and/or its 3 hydroxy derivative and glycine.

Micromonospora colonies usually appear on agar media as pale yellow or light orange colonies, becoming red, brown, blue-green, or purple. Mature colonies become darker, producing brown-black, green-black, or black spores and can take on a mucoid aspect. One of the well defined criteria in the genus *Micromonospora* is the formation of single spores on substrate mycelium. Oddly shaped and multiply septate cells can be found among some enlarged vegetative hyphae.

Micromonospora are well distributed in soil habitats and have been found in fresh and salt water environments.

1.3.4.2 Bioactive metabolites from marine *Micromonospora*

Micromonospora strains isolated from both soil and marine environment have provided many compounds of pharmaceutical interests. The first antibiotic isolated from *Micromonospora* sp. was the aminoglycoside gentamicin (Weinstein *et al.*, 1963). Gentamicin was discovered in 1963 and is produced by both *Micromonospora purpurea* and *Micromonospora echinospora*. It is a broad spectrum antibiotic active against Gram-positive and Gram-negative bacteria.

An increased effort in culturing *Micromonospora* from the marine environment resulted in the isolation of six strains that produce novel bioactive compounds.

Micromonospora sp. strain L31-CLCO-002 and *Micromonospora* sp. strain L-25-ES25-008 were isolated from sponges and discussed in section 1.2.2.1.

Micromonospora lamaivitiensis was isolated from the sea squirt *Polysyncraton lithostrotum* (He *et al.*, 2001). The bacterium produces the anticancer namenamicin as well as the antibiotics lomaiviticin A and B.

Micromonospora marina was isolated from a soft coral collected off the coast of Mozambique (Romero *et al.*, 1997). The strain produces the depsipeptide thiocoraline A which showed antibiotic activities against *S. aureus* (MIC 0.05 μg/ml), *B. subtilis* (MIC 0.05 μg/ml), and *M. luteus* (MIC 0.03 μg/ml). Thiocoraline A is active against P-388, A-549, HT-29 and MEL-28 cell lines.

A *Micromonospora* sp. producing aristotatin A and B, was isolated from sea water collected in Toyama Bay, Japan (Furumai *et al.*, 2000). Aristotatin A and B showed antibacterial activity against Gram-positive bacteria and cytotoxic activity against human myeloid leukemia U937 cell line with an IC50 of 0.4 μg/ml for aristotatin A and 4.0 μg/ml for aristotatin B. Aristotatin A inhibits the neuritogenesis of NGF-stimulated P12 cells at less than 1 μM by inhibiting tubulin polymerization.

1.4 Manzamines

Manzamines are a class of marine alkaloids exclusively found in sponges. The first manzamine, manzamine A, was isolated from a Japanese *Haliclona* sponge in 1986 (Sakai *et al.*, 1986). Manzamines exhibit a complex molecular architecture containing a β-carboline moiety and an unusual polycyclic system. Manzamines have been found in 17 different species representing 5 families of marine sponges all around the world including Okinawa (Sakai *et al.*, 1986), Phillipines (Edrada *et al.*, 1996), Red Sea (Bourguet-Kondracki *et al.*, 1996), Italy (Guo *et al.*, 1998), South Africa (Koren-Goldshlager *et al.*, 1998), Papua New Guinea (Crews *et al.*, 1994) and Indonesia (El Sayed *et al.*, 2001). The bioactivity of manzamines is impressive and diverse. They have shown antitumour, antimicrobial, antiparasitic and insecticidal activities. Though manzamines have had their chemistry and activities thoroughly studied, little is known about the source or the biosynthesis pathway responsible for putting together these complex molecules. Because of the presence of manzamines in multiple sponge species with wide geographical distribution, it has been suggested that manzamine could be of microbial origin but until now, no isolated bacteria or other evidence has supported this hypothesis (Kobayashi & Ishibashi, 1993).

1.4.1 Manzamine A

1.4.1.1 Biogenic pathway

With its complex and intriguing structure and its remarkable pharmaceutical properties, manzamine A (Figure 1.1) has generated a lot of interest regarding its biosynthesis. In 1992, a biogenetic pathway with an intramolecular Diels-Alder reaction was proposed (Baldwin & Whitehead, 1992). Manzamine A is believed to be biosynthesized from an intermediate composed of two dihydropyridine rings with an alkyl residue and a tryptophan unit.
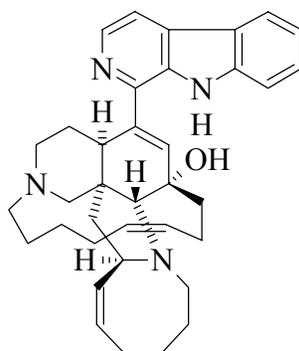
Figure 1.1. Manzamine A

1.4.1.2 Pharmacology of manzamine A

Manzamine A inhibits the growth of P388 mouse leukemia cells (IC$_{50}$ 0.07 µg/ml) (Sakai *et al.*, 1986), and it demonstrated antimicrobial activity against the Gram-positive bacteria *S. aureus* and *B. subtilis* (Nakamura *et al.*, 1987). It shows significant activity

against KB (IC_{50} 0.05 µg/ml), LoVo (IC_{50} 0.15 µg/ml) and HSV-II (MIC 0.05 µg/ml) cells *in vitro* (Ichiba *et al.*, 1994).

Manzamine A has cytoxicity against L1578y mouse lymphoma cells with an ED50 1.8 µg/ml. It has insecticidal activity against the neonate larvae of the polyphagous pest insect *Spodoptera littoralis* (ED_{50} 35 ppm) and an 80% growth inhibition of *S. littoralis* larvae at 132 ppm (Edrada *et al.*, 1996). Manzamine A is also active against *Toxoplasma gondii* and *Mycobacterium tuberculosis* (MIC<12.5 µg/ml) (Ang *et al.*, 2000; El Sayed *et al.*, 2001).

One of the most interesting activities of manzamine A is against malaria. Manzamine A demonstrated activity against *Plasmodium berghei in vivo* by inhibiting more than 90% of the asexual stages after a single intraperitoneal injection into infected mice. Manzamine A is capable of increasing the survival time of highly parasitemic mice and in 40% of the cases, a recovery can be observed after a single injection (Ang *et al.*, 2000). When administrated orally with a mixed oil suspension of manzamine A and 8-hydroxymanzamine A, parasitemia was reduced by 90%. The concentration of manzamine A in the plasma peaked 4 h after injection, remaining high after 48 h. When compared to existing antimalaria drugs chloroquine amd artemisinin, at dosages of 50 and 100 µmol/kg, manzamine A treatment gave significantly better survival times in the treated mice (Ang *et al.*, 2000). Another advantage of manzamine A is the rapid action (1-2 h) against malaria with a sustained high level of drug in the plasma. In addition, manzamine A is effective against strains of *Plasmodium falciparum* that are resistant to other anti-malarial drugs (Hu *et al.*, 2003), suggesting that it is likely to have a novel mode of action.

1.5 Bacterial genomics

Genomics is the sequencing of an organism's genome and the analysis of its gene content. It deals with the systematic use of genome information, associated with other physiological data, to provide answers in biology, medicine, and industry.

Genomics appeared in the 1980s and took off in the 1990s with the initiation of genome projects for several species. Genomics is intimately linked with the method of sequencing DNA which was invented in 1975 by Frederick Sanger, known as the dideoxy termination method (Sanger & Coulson, 1975). Sanger used this method to provide the first entire genome sequence, that of bacteriophage Φ-X174 (Sanger *et al.*, 1978). Since then progress in sequencing technology and larger capabilities permitted the development and completion of large sequencing project such as the Human Genome Project started in 1986. In April 2003, it was announced that 99% of the human genome had been sequenced with 99.9% accuracy.

A total of 403 genomes have been sequenced as of July 2006, with 1596 more genome sequencing projects underway including 932 bacteria, 608 eukaryotes and 56 *Archaea* (http://www.genomesonline.org/).

The generation of billions of base pairs of sequence data created the need for a new field working towards the assembly and analysis of these sequences. The development of bioinformatics fulfilled that need in providing the computational tools required to work with these large amounts of data. Programs like Glimmer can analyze sequence data and detect possible open reading frames (ORFs). Once all possible ORFs have been identified, Glimmer can compare the nucleotide pattern of the ORFs using statistical

model buildt from training set composed of genes of related organisms, to predict which ORFs in the genome are real. The results of the analysis are then used in homology searches where translated ORFs are compared to protein databases, looking for shared sequences. Shared sequences imply shared function but particular attention is paid to homology of binding sites and catalytic sites. Basic Local Alignment Tool (BLAST) (Altschul *et al.*, 1990) can be used to quickly compare the sequence data in hand to a sequence database like GenBank. Different BLAST algorithms are available including BLASTn, that compares two nucleotide sequences and BLASTX that compares translated nucleotide sequence data to the available protein sequence database. BLAST does not match proteins over their entire length but rather look for any regions of similarity within proteins that are closely related. The information collected from the genome sequence analyses are used to deduce which genes are present and what their function may be. Thereafter, each gene-encoding sequence can be sorted according to their assigned role.

1.5.1 Application of genomics to the identification of bioactive compounds

As genomes of organisms have become available, more genes and functions have been discovered, many of them previously unknown. Microbial genomes possess a wealth of genes, many of them unexpressed and/or of unknown function.

The sequencing of *S. coelicolor* A3(2) was a landmark in the field of microbial natural product (Bentley *et al.*, 2002). It provided much information about the organization of

the microbe genome as well as polyketide and non ribosomal peptide biosynthetic gene clusters present that potentially encode bioactive compounds. Bioinformatics tools make it possible to predict products from genes that may not be expressed. Using such approach, coelichelin, a new peptide siderophore, was identified and its structure predicted from *S. coelicolor* genome (Challis & Ravel, 2000).

Genomics has also proved to be a powerful tool in identifying biosynthetic gene clusters from complex microbial communities, this technique being termed environmental genomics or metagenomics. Applying a metagenomic approach to a soil microbial community, two antibiotics, turbomycins A and B were identified after screening more than 24,000 clones (Gillespie *et al.*, 2002). A similar approach but this time targeted at the sponge *D. dissoluta* microbial community, yielded 85 different polyketide synthase (PKS) sequences among which the authors identified a 110 kb PKS sequence encoding for 25,572 amino acids with a predicted molecular mass of ~2.7 MDa (Schirmer *et al.*, 2005). This PKS comprises one starter module and fourteen complete extender modules and one incomplete extender module. The authors predicted that this PKS would be a multimethyl-branched C_{30} fatty acid rather than a complex polyketide.

During the evolution of an organism, the genome is changing, acquiring new capabilities and losing others as its environment changes. Genomics has made it possible to have a comprehensive view of at an organism's genome. Now all genes whether expressed or not can be observed. The world of natural products is expanding with microbes, including some known for a long time, now regarded as an untapped source for new bioactive compounds.

1.5.2 The polyketides

1.5.2.1 General background

Polyketides are a group of secondary metabolites, exhibiting remarkable diversity in their structures and functions. Polyketide natural products are known for their wide range of pharmacologically important activities, including antimicrobial, antifungal, antiparasitic and antitumor properties. These metabolites can be found in many different organisms around the world such as bacteria, fungi, plants, insects, dinoflagellates, mollusks and sponges. Because of their pharmaceutical properties, polyketides have attracted much interest and are among the most sought after natural products. Numerous polyketides have been approved as drugs, such as erythromycin A, a broad spectrum macrolide antibiotic, the antihelmintic agent avermectin and the immunosuppresants FK506 and rapamycin.

1.5.2.2 Synthesis and assembly of polyketides

Polyketides are the result of sequential reactions catalyzed by enzymes called polyketide synthases or PKSs. Polyketide synthases are large multienzyme complexes containing coordinated groups of active sites. Polyketides are synthesized from simple 2, 3 or 4 carbon building blocks such as acetyl-CoA, propionyl-CoA and butyryl-CoA. The essential chain building step of polyketide biosynthesis lies in a decarboxylative condensation analogous to the chain elongation step of classical fatty acid biosynthesis. The difference between fatty acid biosynthesis and polyketide biosynthesis is that for

fatty acids, each successive chain elongation step is followed by a fixed sequence of

ketoreduction, dehydration and enoyl reduction while polyketides, may undergo all, some

or none of these functional group modifications. This ability to undergo these

modification steps combined with the diversity of starting and chain elongation units

results in very diverse molecules with a high level of chemical complexity (Rawlings,

1997).

1.5.2.3   Polyketide synthases

1.5.2.3.1   Type I polyketide synthases

Type I polyketide synthases are very large multifunctional proteins that can be

processive such as those responsible for the synthesis of macrolides like erythromycin

and aureothin (Figure 1.2) or iterative such as those responsible for the synthesis of

lovastatin (Figure 1.3). Modular PKSs represent a unique class of type I polyketide

synthases. Each of these proteins is composed of multiple active domains organized in

modules. Modules are responsible for the creation of a carbon-carbon bond through a

decarboxylative condensation of a ketide extender unit with the growing polyketide

chain, followed by a programmed reductive cycle. In most PKSs, a loading module will

be at the front of module 1 while at the last module, a thioesterase will be unloading the

product.
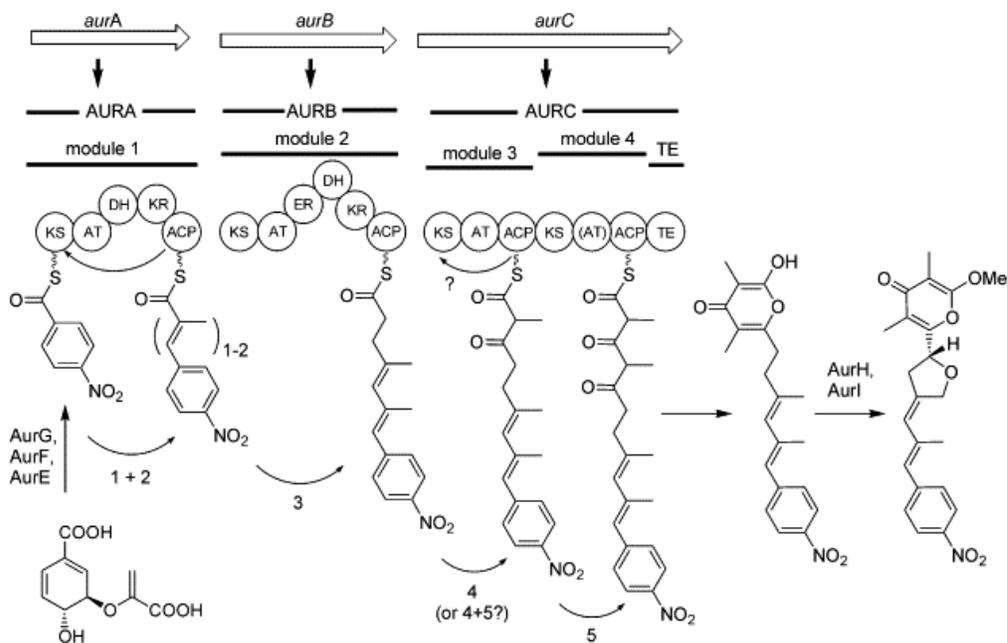
Figure 1.2. Type I processive polyketide synthase, aureothin from *Streptomyces thioluteus* HKI-227. Figure from He and Hertweck (2003).
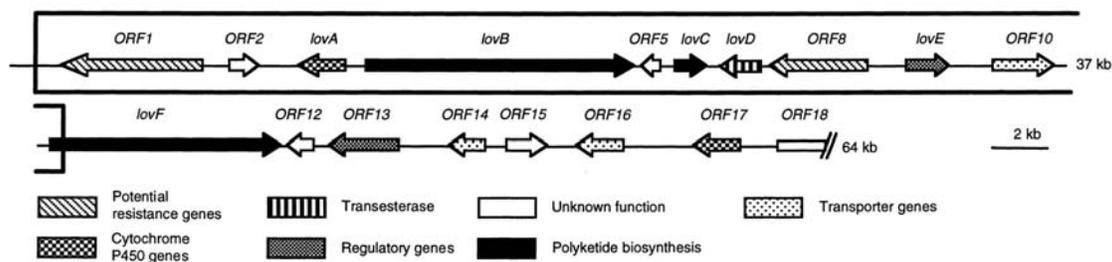
Figure 1.3. Type I iterative polyketide synthase, lovastatin gene cluster. Figure from Kennedy *et al.* (1999).

Iterative type I PKSs are analogous to vertebrate fatty acid synthases. Iterative type I synthases are typically involved in the biosynthesis of fungal polyketides such as 6-methylsaliciclic acid and aflatoxin.

1.5.2.3.1 Type II polyketide synthases

Iterative Type II polyketide synthases (Figure 1.3) are made of complexes of mono-functional proteins. The actives sites of these synthases are distributed among several smaller usually monofunctional polypeptides. Type II polyketide synthases catalyze the formation of molecules that requires aromatization and cyclization without extensive reduction or reduction/dehydratation cycles. These synthases are analogous to bacterial fatty acid synthases. Iterative type II polyketide synthases are responsible for the biosynthesis of actinorhodin in *S. coelicolor* and also in the biosynthesis of tetracenomycin and doxorubicin.
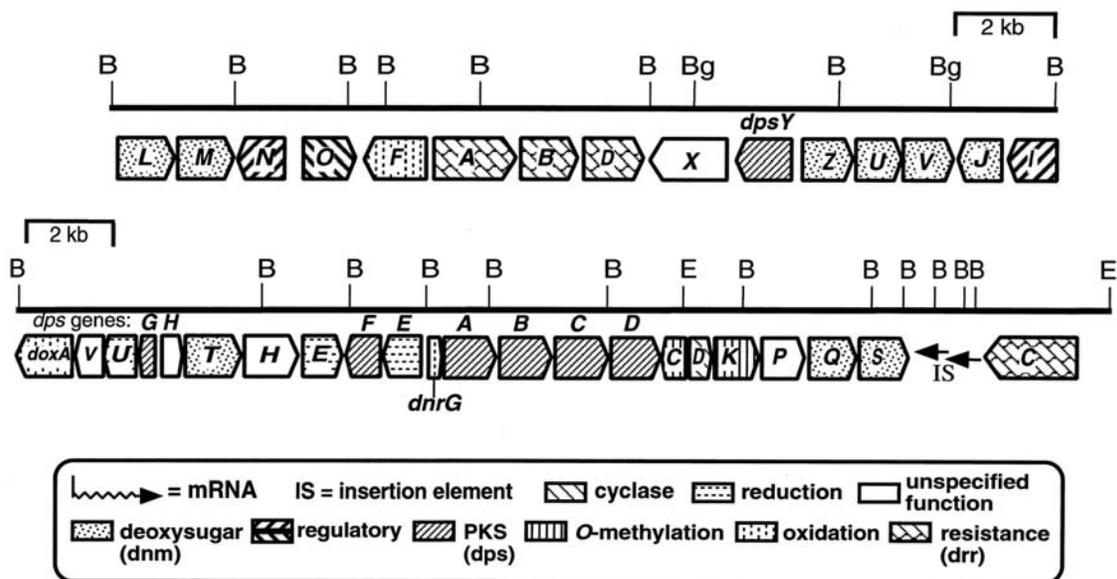
Figure 1.4. Type II iterative polyketide synthase. Physical and functional map of daunorubicin (DNR) and doxorubicin (DXR) gene cluster. Figure from Lomovskaya *et al.* (1999)

1.5.2.3.2   Type III polyketide synthases

Type III PKSs have been found in plants, where they are responsible for the synthesis of chalcones and stilbenes, and in bacteria, where they are responsible for the synthesis of polyhydroxy phenols. Chalcone synthase-like proteins (Figure 1.4) are relatively small proteins when compared to other PKSs. These polyketide synthases are different from the others PKSs as they do not possess a phosphopantetheinyl (P-Pant) arm on which the growing polyketide chains are assembled.
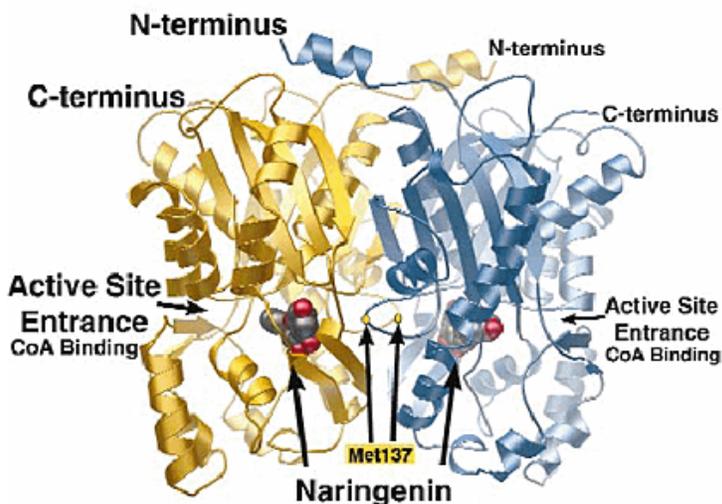
Figure 1.5. Chalcone synthase. Figure from Ferrer *et al.* (1999).

1.5.2.3.3   Domains forming a polyketide synthase

There are seven domains that can compose a polyketide synthase (Figure 1.6).
Acyl-transferases (AT) are responsible for loading the starter, extender and intermediate acyl units, usually acetyl-CoA or malonyl-CoA. Acyl carrier proteins (ACP) hold the

growing chain as a thiol ester. Beta-keto-acyl synthases (KS) catalyze the chain extension.

Beta-keto-reductases (KR) are responsible for the first reduction to an alcohol functionality. Dihydratases (DH) eliminate water to give an unsaturated thiolester. Enoyl reductases (ER) catalyze the final reduction to full saturation. Thiolesterase (TE) is in charge of the macrolide release and its cyclisation (Rawlings, 1997).
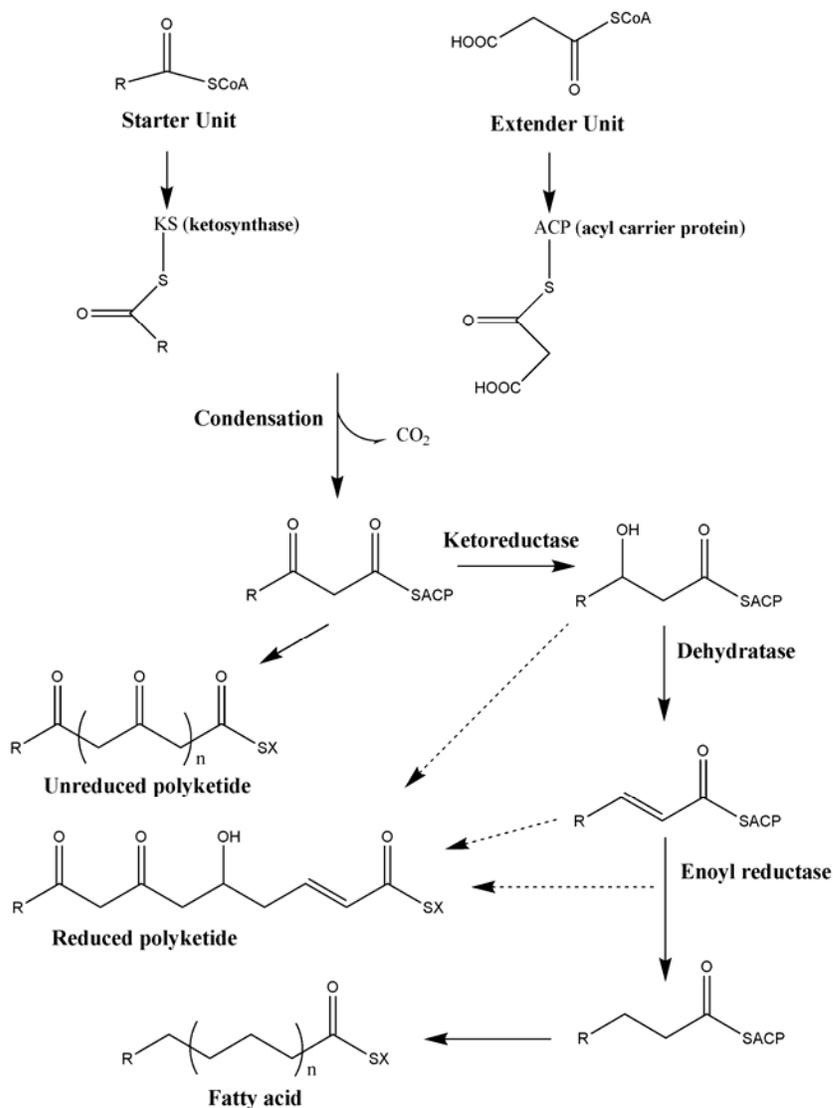
Figure 1.6. Chemistry and role of polyketide synthases various domains. Figure from Gokhale and Tuteja (2001).

1.5.3 Nonribosomal peptides

Nonribosomal peptides are part of a family of complex natural products built from simple amino acid monomers. They can be found in bacteria and fungi where they are synthesized by nonribosomal peptide synthases (NRPS) which are large multimodular, multifunctional proteins. One of the characteristics of NRPS systems is their capability to incorporate proteinogenic as well as non-proteinogenic amino acids in the peptides that are synthesized. Often, these NRPS enzymes work in conjunction with PKS enzymes to generate hybrid products. Nonribosomal peptides as well as the hybrid products are of much interest because of their pharmaceutical properties, i.e. the immunosupressant cyclosporin.

1.5.3.1 Organization of nonribosomal peptide synthases

Nonribosomal peptide synthases are organized in modules (Figure 1.7). Each module is needed for catalyzing one single cycle of a product elongation and modification of the functional group. The number and order of modules as well as the type of domains presents within a module varies and determines the structure of the resulting peptide. An example of a set of domains for the elongation cycle may include a module with adenylation (A), thiolation (T) or peptidyl carrier protein (PCP), a condensation domain (C), an epimerization domain (E) and a thio-esterase (TE). The adenylation domain determines the substrate selection and its covalent fixation on the phospho-pantethein arm of the thiolation domain as a thioester through an AMP-

derivative intermediate. The thiolation domain of peptide synthases, also called peptidyl

carrier protein is the site of 4'-phospho-panthetein (4'-PP) cofactor binding and substrate

acylation. The condensation domain catalizes the formation of the peptide bond between

an aminoacyl or peptidyl-S-PCP from the upstream module and the aminoacyl moiety

attached to the PCP in the corresponding downstream module. It results in the elongation

of the peptide by one residue fixed to the PCP domain in the downstream module. The

epimerization domain epimerizes the innermost amino acid of the peptide chain into the

D-configuration. The thio-esterase module is the last module and is responsible for the

release/cyclization of the completed peptide (Schwarzer *et al.*, 2003).

Nonribosomal peptides often undergo modifications after being synthesized. These

possible modifications include glycosylation, acylation, halogenations, or hydroxylation.

The enzymes catalyzing these reactions are usually associated with the nonribosomal

peptide synthase.
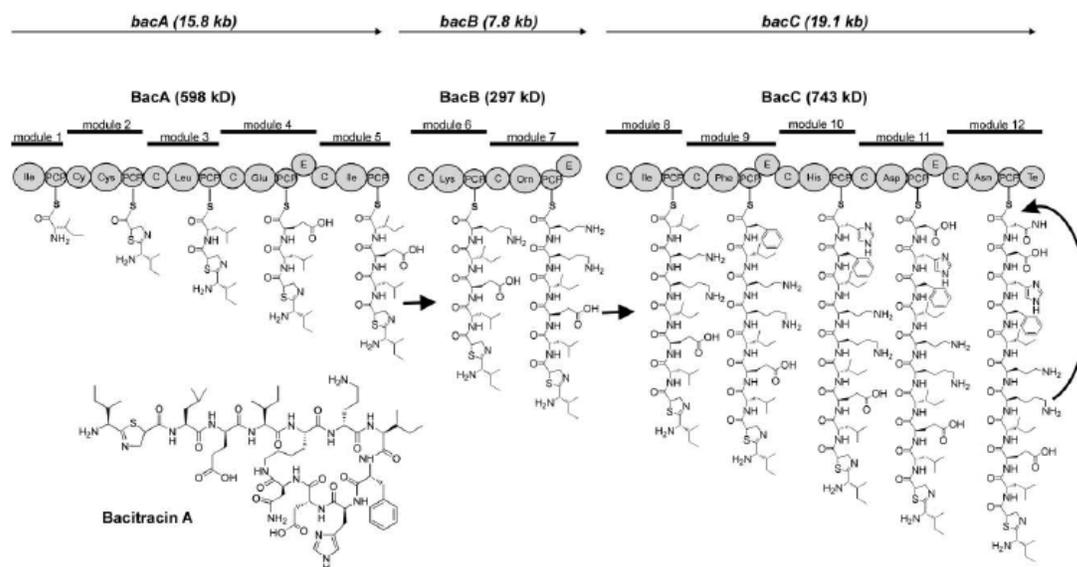
Figure 1.7. The synthesis of bacitracin A requires twelve modules distributed over three NRPSs (BacA, BacB, BacC). Figure from Schwartzer et al. (2003).

1.6 Objectives

This study is part of an ongoing research program focused on the isolation of manzamine A producing bacteria. During the course of this study, a marine sponge was collected off the coast of Manado in Indonesia and sponge-associated bacteria were isolated. Our main research goal was to characterize and cultivate a bacterium capable of producing manzamine A.

The total microbial community associated with the Indonesian sponge *Acanthostrongylophora* was characterized and identified using molecular techniques (**Chapter 2**). We focused on the isolation of sponge-associated bacteria, their phylogenetic identification as well as their ability to produce manzamine A. For that purpose, we used microbiology, chemistry and molecular techniques (**Chapter 3**). Our

discovery of a manzamine-producing actinomycete gave us an opportunity to further characterize this microbe and to attempt to elucidate the manzamine A biosynthetic pathway using molecular and biochemical approaches (**Chapter 4**). A fosmid library of the manzamine-producing bacterium was created and limited sequencing was carried out to give an overview of the genetic capabilities of the microbe (**Chapter 5**), although the biosynthetic pathway for manzamine A was not identified.

The isolation of a manzamine A producing bacterium is especially interesting because of the anti-malarial property of manzamine A. This study may be the first step in providing an alternate anti-malarial drug for which resistance is not yet a concern. It also aimed to overcome for the case of the manzamines, the often limiting supply problem associated with many promising marine drugs.

2 Molecular analysis of *Acanthostrongylophora* sp. microbial communities

2.1 Introduction

Marine sponges have been known for years to contain many associated bacteria that, in some cases, can account for as much as 60 % of the sponge wet weight (Wilkinson, 1978b). Cultivation techniques have long shown their limitations in giving a true representation of what microbial communities really are. The great plate count anomaly (Staley & Konopka, 1985) describes best the discrepancy between the small number of bacteria that can be cultured on agar plates compared to the large number of uncultivable ones that are enumerated by microscopic counting.

The development of gene cloning and 16S rRNA gene sequencing techniques has permitted a better description of the complex and diverse microbial communities found in many environmental samples. The advantage of such techniques is that they are not dependent on the cultivation of microbes. The analysis of the results of 16S rRNA gene clone libraries gives a more complete view of the bacterial composition than culture-based studies. The phylogenetic analysis of 16S rRNA gene sequences permits putative identification of bacteria by comparison with sequences from known closely related organisms. This information can be used to design media containing specific nutrients or vitamins in order to increase the chances of growing organisms of interest (Webster & Hill, 2001).

Molecular analysis of microbial diversity is a technique that gives a good overview of the diversity of a microbial community. However, one of the limitations lies in the use of PCR prior to the cloning step. PCR is a very sensitive technique that is used to catalogue bacterial and archaeal species richness in the environment. The products of the amplification of a mixture of 16S rRNA genes from natural communities are considered representative of the original microbial community. Limitations for PCR exist and should be taken into account when analyzing the results. A study showed that PCR bias can be caused by template annealing in the amplification of the pool of 16S rRNA genes targeted for amplification by PCR (Suzuki & Giovannoni, 1996). Factors that can contribute to this bias are the guanosine plus cytosine (G+C) content and the length of template DNA. There is a selective amplification of template with low G+C content as more energy is required to denature high G+C genes into single stranded molecule (Suzuki & Giovannoni, 1996). This may lead in certain cases to an over-representation of low G+C genes in the clone library. PCR bias can be reduced by applying high ramp rates between the denaturation and annealing steps, low annealing temperatures should be used and extension times >180 s should be avoided (Acinas *et al.*, 2005). The formation of secondary structures of the template can limit the access to the priming site and hamper the amplification (Meyerhans *et al.*, 1990), representing another potential source of bias. These factors affect the amplification efficiency and can lead to a product concentration not representing the original template concentration. The fidelity with which the DNA polymerase *Taq* replicates the DNA templates can also be a source of bias but can be limited by using high fidelity DNA polymerase less prone to incorporation errors (Becker *et al.*, 2000). During the amplification of rRNA genes from

mixed microbial populations the formation of chimera can occur. A chimera forms when a prematurely terminated amplicon reanneals to a foreign DNA strand and is replicated in the subsequent PCR cycles (Hugenholtz & Huber, 2003). This results in a single sequence composed of two or more phylogenetically distinct parent sequences. Chimeric sequences can be detected by subjecting clone library sequences to programs such as Check_Chimera (Maidak *et al.*, 1999) or Bellerophon (Huber *et al.*, 2004). Reducing the number of cycles from 30 to 25 or 20 cycles (e.g., until a band is barely visible on agarose gels) can reduce both the chimeras and *Taq* DNA polymerase errors (Acinas *et al.*, 2005).

Molecular analysis of a microbial community indexes the different bacteria present. Though it does not provide reliable quantitative information, it can be used to give a rough estimate of the abundance of a particular organism or group of organisms within the microbial community. Even taking into account the PCR limitations, it can be assumed that the abundance of sequences in clone libraries approximately reflects that of organisms in their environment. Denaturant gradient gel electrophoresis (DGGE) is another molecular technique used to assess the complexity of microbial communities. The presence of bacterial groups highlighted in the clone library can be verified by microscope observations using fluorescent *in situ* hybridization (FISH) probing which allows specific cells to be visualized and enumerated (Friedrich *et al.*, 1999; Webster *et al.*, 2001b).

The relationship between marine sponges and their associated bacteria has generated much interest. However, most of the roles ascribed to sponge symbionts remain putative. Sponge-associated bacteria have been proposed to contribute to nutrient acquisition

(Borowitzka *et al.*, 1988), stabilization of the sponge skeleton (Wilkinson *et al.*, 1981), processing of metabolic waste (Wilkinson, 1978c), protection from UV light (Shick & Dunlap, 2002) and chemical defense (Schmidt *et al.*, 2000).

Indonesian sponges *Acanthostrongylophora* sp. Sponge 35 and Sponge 52 were studied because manzamines were detected in these sponges. It has been hypothesized that manzamines could be of microbial origin (Kobayashi & Ishibashi, 1993). The microbial community analysis of these two manzamine-containing sponges, *Acanthostrongylophora* sp. Sponge 35 and Sponge 52, and their comparison may provide information related to the suspected microbial origin of manzamines.

2.2 Materials and Methods

2.2.1 Sampling

Specimens of the marine sponges *Acanthostrongylophora* sp. Sponge 35 and Sponge 52 were collected by SCUBA diving from reef slopes and vertical surfaces of Black Reef Point (Manado Bay, Indonesia) at depth between 6 and 33 m. Sponges were transferred directly to plastic bags containing seawater to prevent contact of sponge tissue with air. In the laboratory, sponge tissue was kept at -80˚C until used. The sponge tissue was then freeze-dried prior to molecular manipulation. Water samples were collected in sterile 20-liter containers adjacent to the sponges sampled. Water samples (10 liters) were filtered through 0.22 µm-pore-size Sterivex filters (Millipore).

2.2.2 Sponge DNA extraction

DNA was extracted from freeze-dried sponge tissue using a bead-beater method adapted from Pitcher *et al.* (1989). Dried tissue (4 g) was ground with a mortar and pestle and resuspended in 16 ml of TE buffer (10 mM Tris and 1 mM EDTA, pH 8.0) and 4 ml of isoamyl alcohol. The solution was transferred to a 50 ml bead-beater chamber (Biospec Products, Inc, Bartlesville, OK). Zirconia/Silica beads (0.1 mm and 1 mm) were added to 1/4 volume of the chamber. The bead-beater was run 4 times for 1 min. The solution was transferred in a 50 ml tube, 10 ml of guanidium thiocyanate buffer was added, mixed gently, and transferred to ice. Ammonium acetate (10 M) was added to a 2.5 M final concentration. Standard phenol/chloroform extraction and chloroform/isoamyl alcohol extraction (Appendix 1) followed. DNA was precipitated with cold isopropanol, cleaned with 70% ethanol, and resuspended in TE buffer (pH 8). DNA was quantified using a spectrophotometer.

2.2.3 PCR amplification and cloning

For bacteria, PCR was performed using 100 ng of DNA with universal 16S rRNA gene primers 8-27f (Weisburg *et al.*, 1991) and 1492r (Reysenbach *et al.*, 1992). The DNA amplification was carried out using Hi-Fi Platinum Taq (Invitrogen, San Diego, Calif.). Cycling conditions were as follow: initial denaturation at 94ºC for 5 min, 20 cycles of 94ºC for 30 sec, 48ºC for 2 min, 72ºC for 1.5 min, and a final extension of 5

min at 72ºC in a PTC-200 MJ-research thermal cycler (Bio-Rad, Hercules, Calif). For

archaea, PCR was performed using 100 ng of 16S rRNA gene primers 21f and 951r

(DeLong, 1992). Cycling conditions were: initial denaturation at 94ºC for 5 min, 25

cycles of 94˚C for 1 min, 54˚C for 30 sec, 72˚C for 2 min, and a final extension of 5 min

at 72˚C. PCR products were purified by electrophoresis in a 1% (wt/vol) agarose gel and

bands of approximately 1500 bp were excised and recovered using a gel extraction kit

(Qiagen, Inc., Chatsworth, Calif.). Purified PCR products were cloned with a TOPO-XL

cloning kit according to the manufacturer's instructions (Invitrogen).

2.2.4 Sequencing, phylogenetic and rarefaction analysis

Sequencing of clones was done using an ABI 377 automated sequencer (PE

Applied Biosystems, Foster City, Calif.) with M13 forward and reverse primers.

Sequences were compared to those in the GenBank database with the Basic Local

Alignment Search Tool (BLAST) algorithm to identify known closely related sequences.

Sequences were examined for the formation of chimeras using the program

CHECK_CHIMERA (Maidak *et al.*, 1997). Sequences of 142 clones, with a length of at

least 500 bp, were manually compiled and aligned with Phydit software (Chun, 1995).

Trees were generated by the neighbor-joining (Saitou & Nei, 1987) algorithm (Kluge &

Farris, 1969) implemented in Phydit. The robustness of inferred tree topologies was

evaluated after 1,000 bootstrap resamplings of the neighbor-joining data and only values

>50% were shown.

The assemblage of 16S rRNA gene sequences in each library was analyzed by rarefaction analysis using EcoSim (Gotelli & Entsminger, 2006) to assess the extent to which the diversity of microbial communities was represented by the library at the class and species level. The number of species in each clone library was determined by comparing closely related sequences using bl2seq (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi) (Tatusova & Madden, 1999). 16S rRNA equences exhibiting a percentage of similarity of 97% or lower (Devereux *et al.*, 1990) were considered two different species.

2.2.5 Fluorescence *in situ* hybridization

Sponge tissue was fixed in 4% paraformaldehyde for 8 h at 4˚C and kept in 70% ethanol until processed. Sections of sponge tissue (ca. 1 cm^3) were first immersed in a 15% sucrose solution for 3 h and then transferred in a 15% sucrose: OCT resin (Tissue-TEK, Sakura Finetek Inc., Torrance, Calif) series of 3 sucrose: 10 OCT, 1 sucrose: 10 OCT, 1 sucrose: 30 OCT and 100% OCT for a minimum of 2 h at each step. OCT moulds (Tissue-TEK cryomould, Elkart, In.) were placed on dry ice and a small quantity of OCT allowed to freeze in the bottom of the moulds prior to the addition of samples. Samples in molds were set on dry ice and subsequently stored at -20˚C. Tissue sections were cut to a thickness of 20 to 30 μm using a cryomicrotome (Leitz 1720C) and transferred to silane coated slides (2% aminosilane: 3-aminopropyltriethoxy-silane, Sigma, St. Louis, MO). Prior to hybridization, sections were photo-bleached under a mercury lamp for 60 sec to reduce the sponge tissue autofluorescence. All probes were labeled with the indocarocyanine fluorochrome Cy3 and were synthesized by Sigma-

Genosys. Oligonucleotide probes used in this study are listed in Table 2.1.

Hybridization solution is made of 360 µl of 5 M NaCl, 40 µl of Tris-HCl (pH 7.2), 0.01%

of 10% SDS, formamide (concentration between 20% and 35% according to the probe

used), H_2O added to a final volume of 2 ml (Table 2.2). Eight µl of the hybridization

solution was mixed with 1 µl (25 ng) of the appropriate fluorescent probe and applied to

each section of the sponge tissue. Samples were incubated at 46˚C for 3 h in a wet

chamber inside the slide chamber Unit Alpha (Bio-Rad, San Francisco, Calif) adapted on

the PTC-200 MJ thermal cycler (Bio-Rad). After hybridization, slides were carefully

removed and rinsed immediately in prewarmed wash buffer. The wash buffer contains

2.15 ml of 5 M NaCl, 1 ml of 1 M Tris-HCl (pH 7.2), 50 µl of 10% SDS, formamide

(concentration between 20% and 35% according to the probe used), H_2O to 50 ml (Table

2.3).

Table 2.1. Sequence of oligonucleotide probes used for FISH

Probe	Sequence (5' to 3')	16S rRNA target site[a]	Specificity	Formamide[b]	Reference
Eub338	GCTGCCTCCCGTAGGAGT	338-355	Bacteria	20 %	Amann *et al*, 1990
NonEub338	ACTCCTACGGGAGGCAG	338-355	Bacteria	20 %	Manz *et al*, 1992
HGC69a	TATAGTTACCACCGCCGT	1901-1918	Actinobacteria	25 %	Roller *et al*, 1994
GAM42	GCCTTCCCACATCGTT	1027-1043	γ-protebacteria	35 %	Amann *et al*, 1990

[a] rRNA E.coli numbering; [b] All hybridization temperatures were 46˚C

Table 2.2. FISH hybridization buffer composition

Probe	Formamide %	H$_2$O	SDS 10%	Tris 1M
Eub338	20%	1198 µl	2 µl	40 µl
NonEub338	20%	1198 µl	2 µl	40 µl
HGC 69A	25%	1098 µl	2 µl	40 µl
Gam 42	35%	998 µl	2 µl	40 µl

Table 2.3. FISH washing buffer composition

Probe	Formamide %	H$_2$O	SDS 10%	Tris 1M	5 M NaCl
Eub338	20%	up to 50 ml	50 µl	1 ml	2150 µl
NonEub338	20%	up to 50 ml	50 µl l	1 ml	2150 µl
HGC 69a	25%	up to 50 ml	50 µl	1 ml	1490 µl
Gam 42	35%	up to 50 ml	50 µl l	1 ml	700 µl

2.2.6 Denaturing gradient gel electrophoresis (DGGE) analysis

DNA from seawater samples was extracted by the method of Somerville *et al.* (1989). Total DNA from *Acanthostrongylophora* sp. Sponge 35, Sponge 52 and seawater were analyzed by DGGE. Primers P2 and P3 (Muyzer *et al.*, 1993) were used to amplify the 194 bp region corresponding to positions 341 and 534 in the 16S rRNA gene of *E. coli*. PCR amplification was performed on 100 ng of DNA with Platinum *Taq* (Invitrogen Life Technologies) and 25 pmol of each primer. The cycling conditions were a 5 min hot start at 94°C; 30 cycles of 1 min at 92°C, 1 min at 55°C, and 1 min at 72°C; and a final 5 min extension step at 72°C. Thermal cycling was performed in a PTC-200

cycling system (MJ Research, Research, Waltham, Mass.). The final PCR product was loaded onto a 6% acrylamide gel with a denaturing gradient of 45 to 65%. Electrophoresis was performed using the D-Code system (Bio-Rad, Calif.) in 1x TAE (20 mM Tris acetate, 10 mM sodium acetate, 0.5 mM EDTA) at a constant temperature of 60°C and voltage of 60 V for 16 h. The gel was stained with 1x SYBR green (Molecular Probes Inc., Eugene, Oreg.) for 10 min and visualized with the Typhoon 9410 image system (Amersham Biosciences, United Kingdom).

2.3 Results

2.3.1 Bacterial diversity of two Indonesian sponges indicated by molecular techniques

2.3.1.1 Diversity of bacteria in sponge *Acanthostrongylophora* sp. Sponge 35 and Sponge 52 indicated by 16S rRNA gene clone library analysis

PCR performed on the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 and Sponge 52 yielded a product of the expected size of 1450 bp. The PCR products were cloned using TOPO-XL vector (Invitrogen) and yielded 250 clones for Sponge 35 and 200 for Sponge 52 of which 192 for each were selected. For Sponge 35, 142 clones were sequenced and 108 non-redundant clones with at least 500 bp were subjected to phylogenetic analysis. Forty one clones of Sponge 52 clone library were sequenced and phylogeneticaly analyzed.

The sequence data indicate that a high diversity of bacteria phylotypes was present within Sponge 35 (Figure 2.1). Overall, 22% of the clones clustered with *Acidobacteria*, 21% with *Actinobacteria*, 19% with *Chloroflexi*, 12% with alpha-proteobacteria, 11% with gamma-proteobacteria, 6% with delta-proteobacteria, 6% with *Bacteroidetes* and 3% with spirochaetes.

The analysis of the forty one Sponge 52 clones (Figure 2.2) showed that 39% of the clones clustered with gamma-proteobacteria, 27% with *Actinobacteria*, 12% with *Acidobacteria*, 7% with nitrospira, 5% with planctomycetes, 5% with cyanobacteria, 3% with *Chloroflexi* and 3% with alpha-proteobacteria.

Rarefaction analyses were conducted at the bacterial class and species level on both Sponge 35 and 52 clone libraries. For *Acanthostrongylophora* sp. Sponge 35 16S rRNA gene clone library, the analysis showed that it was appropriately sampled at the class level reflecting well the sponge microbial group diversity (Figure 2.3). At the species level, the rarefaction analysis suggests that the library is slightly undersampled (Figure 2.4). For *Acanthostrongylophora* sp. Sponge 52, the rarefaction analysis showed that the clone library was undersampled both at the class (Figure 2.5) and species level (Figure 2.6).

Analyzed sequences of clones from both Sponge 35 and 52 clone libraries did not match exactly to previously known bacterial species. In 80% of the cases, the 16S rRNA gene sequences from clones were most closely related to previously described sponge-associated bacteria, based on BLAST analysis.

Six phylogenetic subgroups were created for the analysis of the 16S rRNA gene sequences from the clones (Table 2.4).

Table 2.4. Description of *Acanthostrongylophora* Sponge 35 and Sponge 52 subgroups

Subgroup	Closely related bacterial class
I	*Actinobacteria*
II	alpha- and gamma-proteobacteria
III	*Chloroflexi* and *Spirochaetes*
IV	*Acidobacteria*
V	delta-proteobacteria and *Bacteroidetes*
VI	*Nitrospira, Planctomycetes* and *Cyanobacteria*
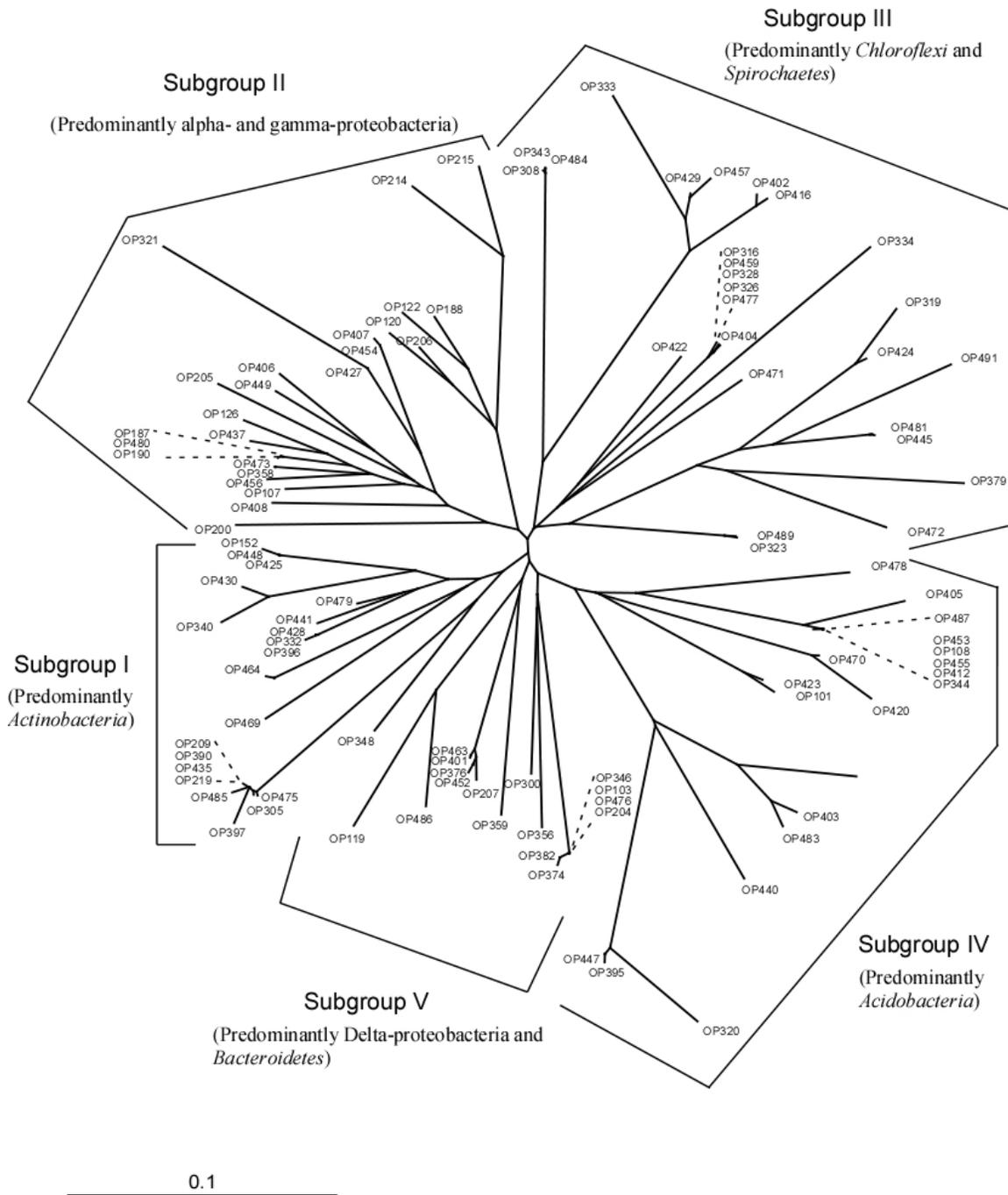
Figure 2.1. Radial cladogram showing the diversity of bacterial clone sequences from the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35. The neighbor-joining tree is based on 500 bp of 16S rRNA gene sequence and includes the 108 clones for which unique sequence was obtained. The scale bar represents 0.1 substitutions per nucleotide position. Dotted lines indicate the original position of very closely related clones.
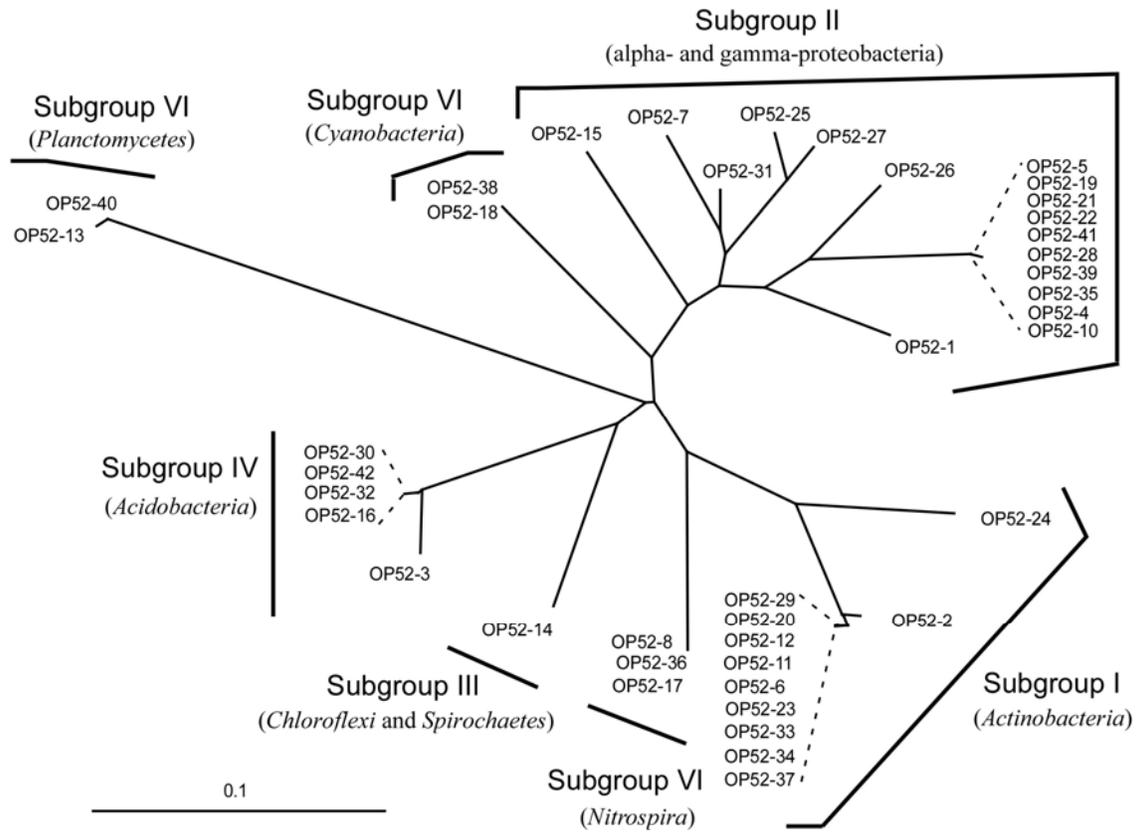
Figure 2.2. Radial cladogram showing the diversity of bacterial clone sequences from the Indonesian sponge *Acanthostrongylophora* sp. Sponge 52. The neighbor-joining tree is based on 500 bp of 16S rRNA gene sequence and includes the 41 clones for which unique sequence was obtained. The scale bar represents 0.1 substitutions per nucleotide position. Dotted lines indicate the original position of very closely related clones.

69
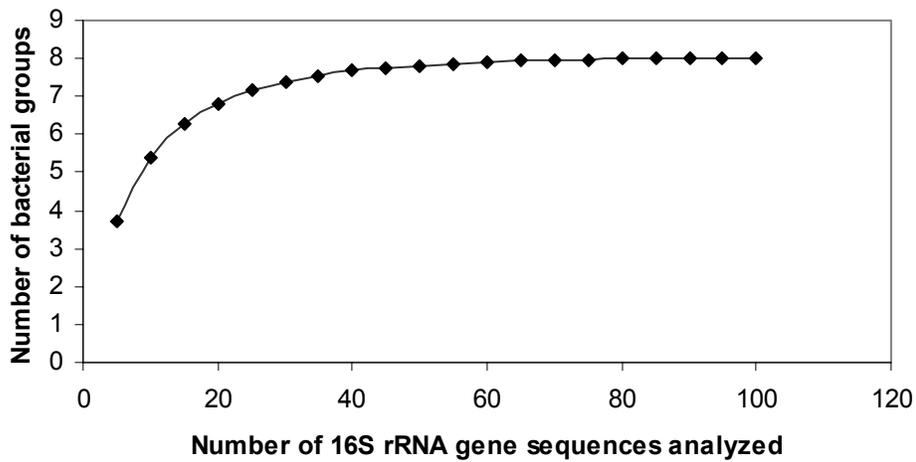
Figure 2.3. Rarefaction analysis of the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 showing the bacterial diversity at the class level, relative to the number of 16S rRNA gene sequences analyzed.
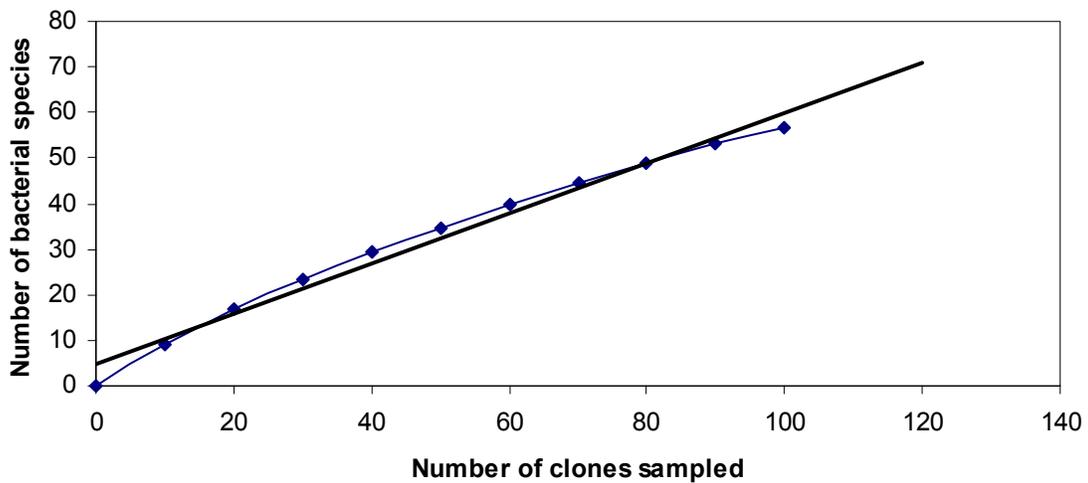
Figure 2.4. Rarefaction analysis of the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 showing the bacterial diversity at the specie level, relative to the number of 16S rRNA gene sequences analyzed. In blue is the actual number of species and in black is the projected number of species according to the sample size.
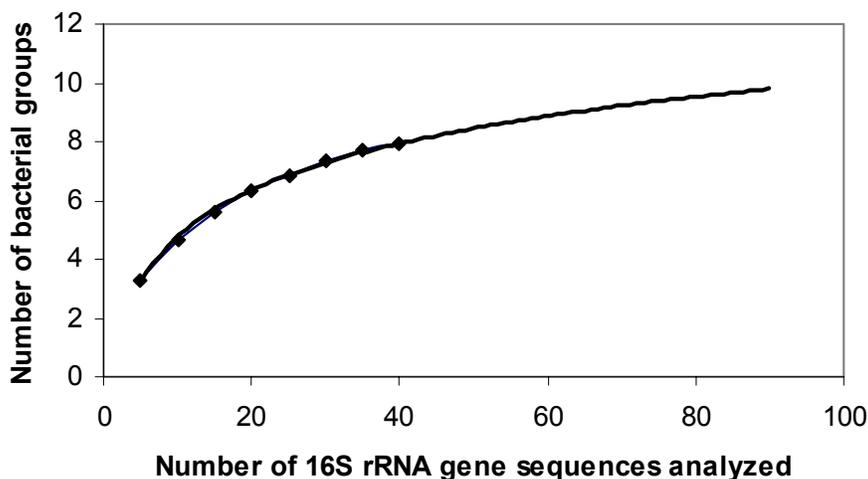
Figure 2.5. Rarefaction analysis of the Indonesian *Acanthostrongylophora* sp. Sponge 52 showing the bacterial class diversity relative to the number of 16S rRNA gene sequences analyzed. Diamonds indicate the actual sequences analysed and in black is a regression curve showing the potential diversity as the size of sample increases.
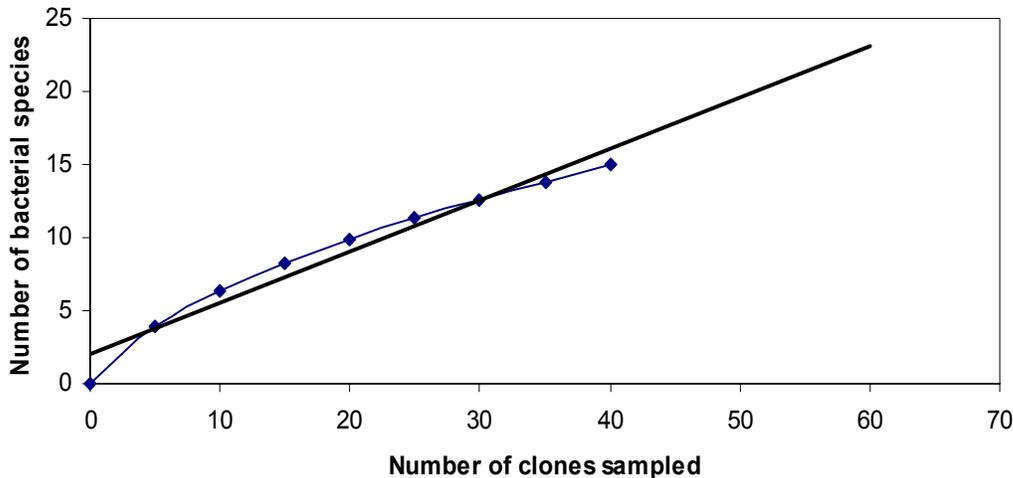
Figure 2.6. Rarefaction analysis of the Indonesian *Acanthostrongylophora* sp. Sponge 52. showing the bacterial species diversity, relative to the number of 16S rRNA gene sequences analyzed. In blue is the actual number of species and in black is the projected number of species according to the sample size.

Clones from *Acanthostrongylophora* Sponge 35 are numbered OPXXX and clones from

Acanthostrongylophora Sponge 52 are numbered OP52-XXX. Details of close relatives

of clones based on BLAST analysis in each subgroup are as follows:

Subgroup I

Subgroup I is composed of *Actinobacteria*-related clones and can be broken down in five

main groups (Figure 2.5). The first one comprises eight clones and is closely related to

Theonella sponge symbiont JAWS10 (AF434968) (Hentschel *et al.*, 2002).

The second group comprises five clones. Clones OP396, OP332 and OP428 are closely

related to *Theonella* sponge symbiont RSWS15 (AF434943) (Hentschel *et al.*, 2002).

OP479 is closely related to *Theonella* sponge symbiont JAWS11 (AF434969) and OP441

is closely related to a deep-sea octocoral symbiont ctg_CGOAA22 (DQ395502).

The third group is composed of three clones, OP340, OP430 and OP52-24 that are all

closely related to *Theonella* sponge symbiont RSWS10.

The last group is composed of thirteen clones from which ten are from Sponge 52. Their

closest relatives is the Aplysina sponge symbiont TK99 (AJ347072) (Hentschel *et al.*,

2002).

Subgroup II

Subgroup II is composed of 13 alpha-proteobacteria and 28 gamma-proteobacteria

(Figure 2.6). Most representatives of the alpha-proteobacteria group are closely related to

other uncultured sponge-associated bacteria obtained from a variety of sponges including

Theonella sp. (OP454, 407, 408, 321, 427, OP52-15) (Hentschel *et al.*, 2002),

Cymbastela concentrica (OP120,122, 188, 206, 214, 215) and Dictyoceratid sponges (OP207) (Ridley *et al.*, 2005b). One clone, OP200, is closely related to the uncultured salt marsh bacterium SIMO-992 (AY712437).

The gamma-proteobacteria group also has clones related to a symbiont of the bivalve *Codakia orbicularis* (OP107) (Durand *et al.*, 1996), uncultured sea-floor bacterium (OP406, 449) (DeLong *et al.*, 2006), bacteria associated with the coral *Oculina patagonica* (OP205, 358, 473) and sponge symbionts (OP205, 126, 437, 456, 190, 480, 187) (Enticknap *et al.*, 2004; Hentschel *et al.*, 2002). A group of ten clones, all belonging to Sponge 52 library, is clustering together and is closely related to *Stenotrophomonas maltophilia* (DQ077704), a bacterium know to degrade the herbicide thifensulfuron-methyl.

Subgroup III

Subgroup III is composed of eight spirochaetes and 19 *Chloroflexi*-related bacteria (Figure 2.7). Five spirochaetes *clones*, OP457, 33, 429, 402 and 416 are related to *Aplysina* sponge symbiont TK79 (AJ347081) (Hentschel *et al.*, 2002). The other three, OP484, 343 and 308 are closely related to the uncultured marine sediment clone ODP1251B15 (AB177319) (Inagaki *et al.*, 2006). Of the 19 *Chloroflexi* clones, only two are not closely related to previously described sponge-associated bacteria. Clones OP481 and OP445 are closely related to the uncultured *Chloroflexi* bacterium SAR269 (AY534090) from the Sargasso Sea (Morris *et al.*, 2004). A distinct clade supported by a bootstrap value of 100 included six clones, OP404, 326, 477, 328, 316, 459 and was closely related to the *Aplysina* symbiont TK63 (AJ347076) (Hentschel *et al.*, 2002).

Subgroup IV

Subgroup IV is composed of *Acidobacteria* realated sequences (Figure 2.8). The closest

relatives for all *Acidobacteria* clones are previously identified sponge symbionts of

Theonella swinhoei (Hentschel *et al.*, 2002) or *Discodermia dissoluta* (Schirmer *et al.*,

2005). A group composed of seven clones, two from Sponge 35 and five from Sponge

52, formed a tight cluster closely related to *Theonella* sponge symbiont PAUC54

(AF186439) (Hentschel *et al.*, 2002). Another clade, supported by a bootstrap value of

100 was formed by eight clones all from Sponge 35 library and closely related to

Theonella sponge symbiont PAUC37 (AF186413) (Hentschel *et al.*, 2002).

Subgroup V

Subgroup V is composed of six *Bacteroidetes* and seven delta-proteobacteria, all from

Sponge 35 clone library (Figure 2.9). *Bacteroidetes*-related clones form one clade

closely related to *Theonella* sponge symbiont JAWS2 (AF439975) (Hentschel *et al.*,

2002). Delta-proteobacteria-related clones form two clades, one with two clones, OP486

and OP119, closely related to the isolate *Bacteriovorax* sp. strain ETC (AY294217)

(Davidov & Jurkevitch, 2004) and the uncultured costal bacterioplankton bacterium

PI_4t11d (AY580409) (Acinas *et al.*, 2004). The second clade is composed of the

remaining five clones and is closely related to the Pacific Ocean sediment clone D132

(AY375137) (Zeng *et al.*, 2005).

Subgroup VI

Subgroup VI is composed of two *Planctomycetes,* two *Cyanobacteria* and three *Nitrospira*, all originating from Sponge 52 clone library (Figure 2.10). Of the *Planctomycetes,* OP52-40 is closely related to the uncultured marine eubacterium HstpL83 (AF159642) that was associated with the seagrass *Halophila stipulacea* and OP52-13, is closely related to the coral symbiont CD207D05 (DQ200597). The two *Cyanobacteria* clones, OP52-18 and OP52-38 are closely related to *Synechococcus* sp. strain WH8012 (AF539812), a phototrophic picoplankton. The three *Nitrospira* clones, OP52-8, OP52-17 and OP52-36 are all closely related to *Theonella* sponge symbiont JACN1 (AF434964) (Hentschel *et al.,* 2002).
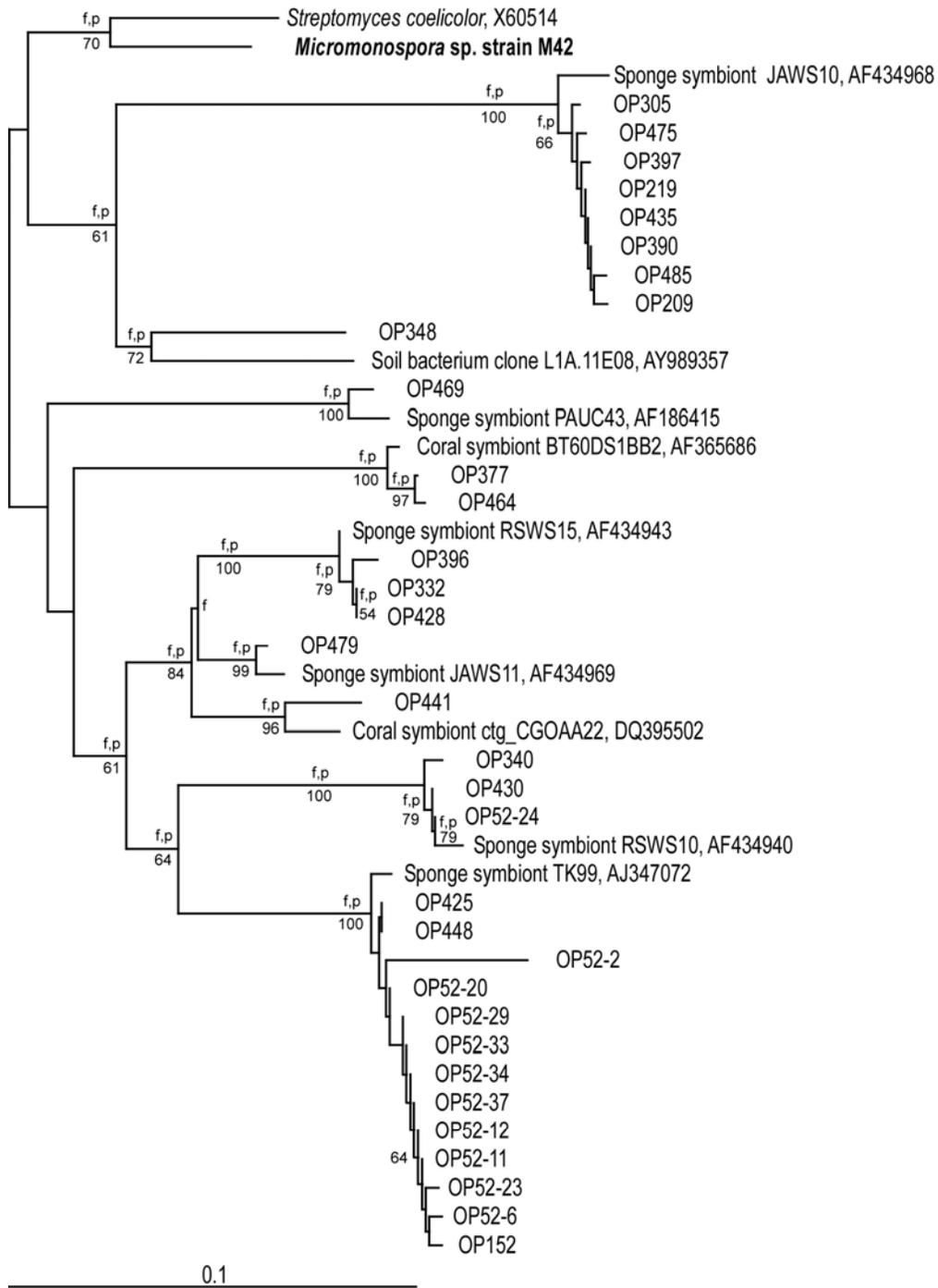
Figure 2.7. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with the *Actinobacteria* (Subgroup I). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Only values > 50% are shown. Scale bar represents 0.1 substitutions per nucleotide position. *E. coli* was used as the outgroup.
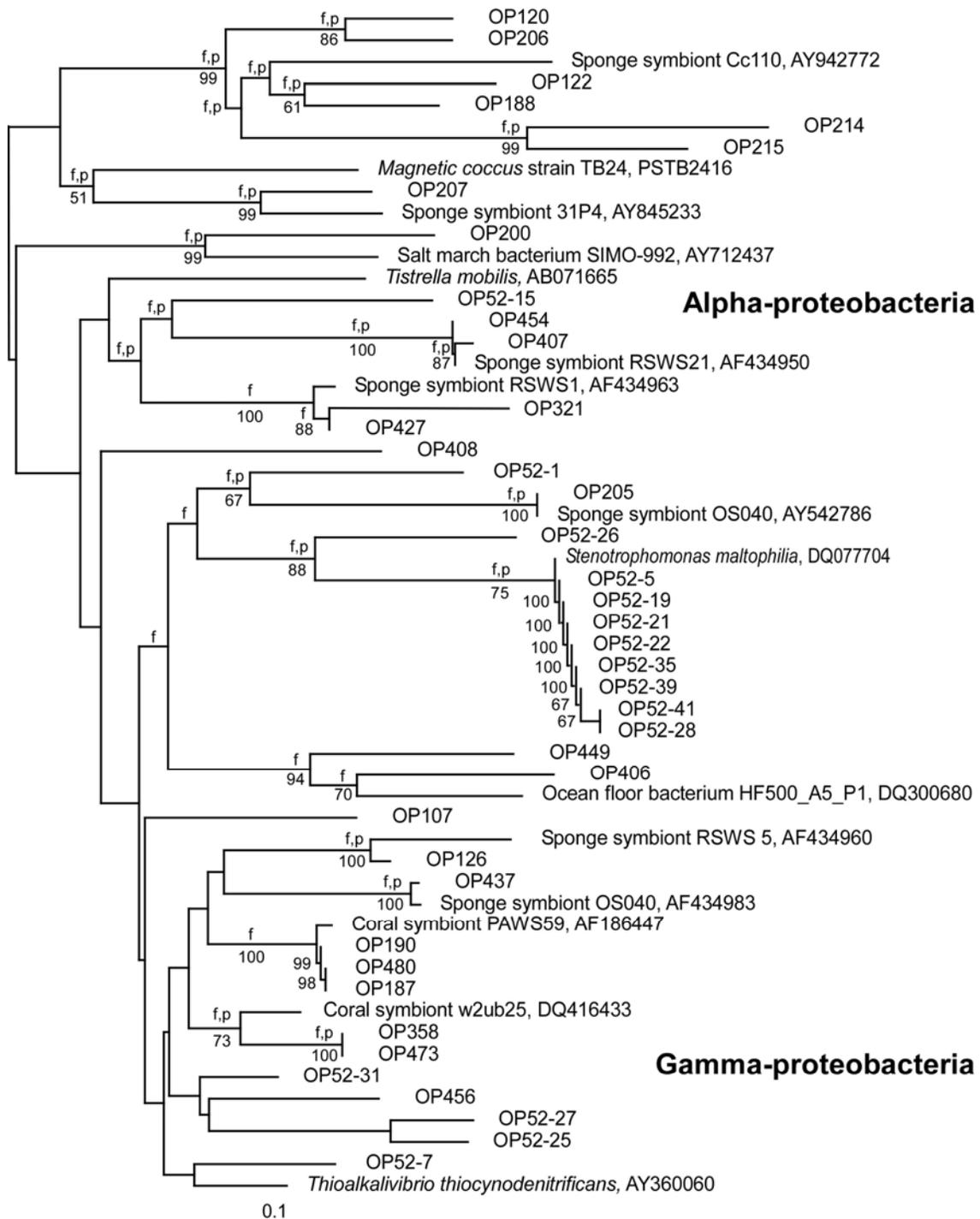
Figure 2.8. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with alpha- and gamma-proteobacteria (Subgroup II). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position. *S. coelicolor* (X60514) was used as the outgroup.
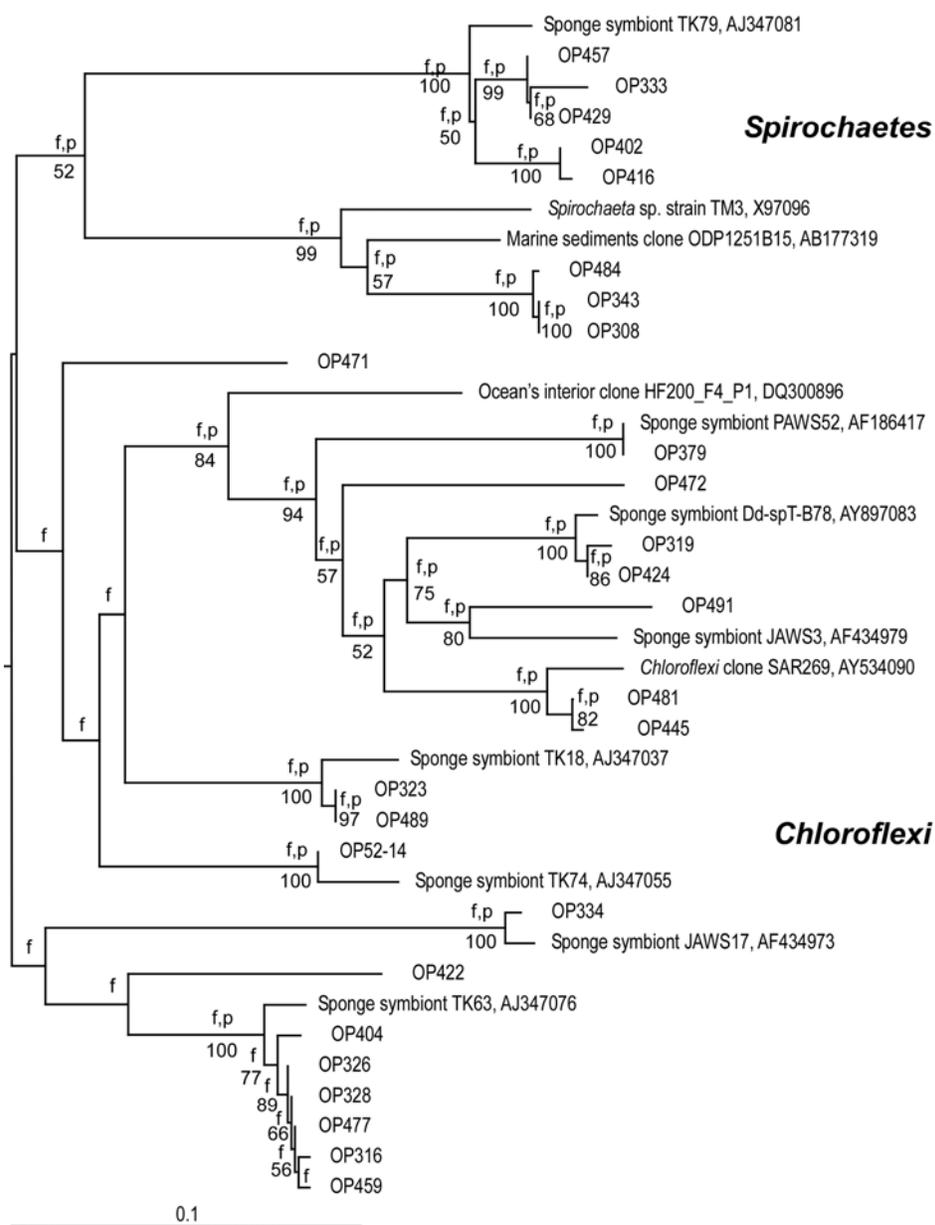
Figure 2.9. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with the *Chloroflexi* and *Spirochaetes* (Subgroup III). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position. *E. coli* was used as the outgroup.
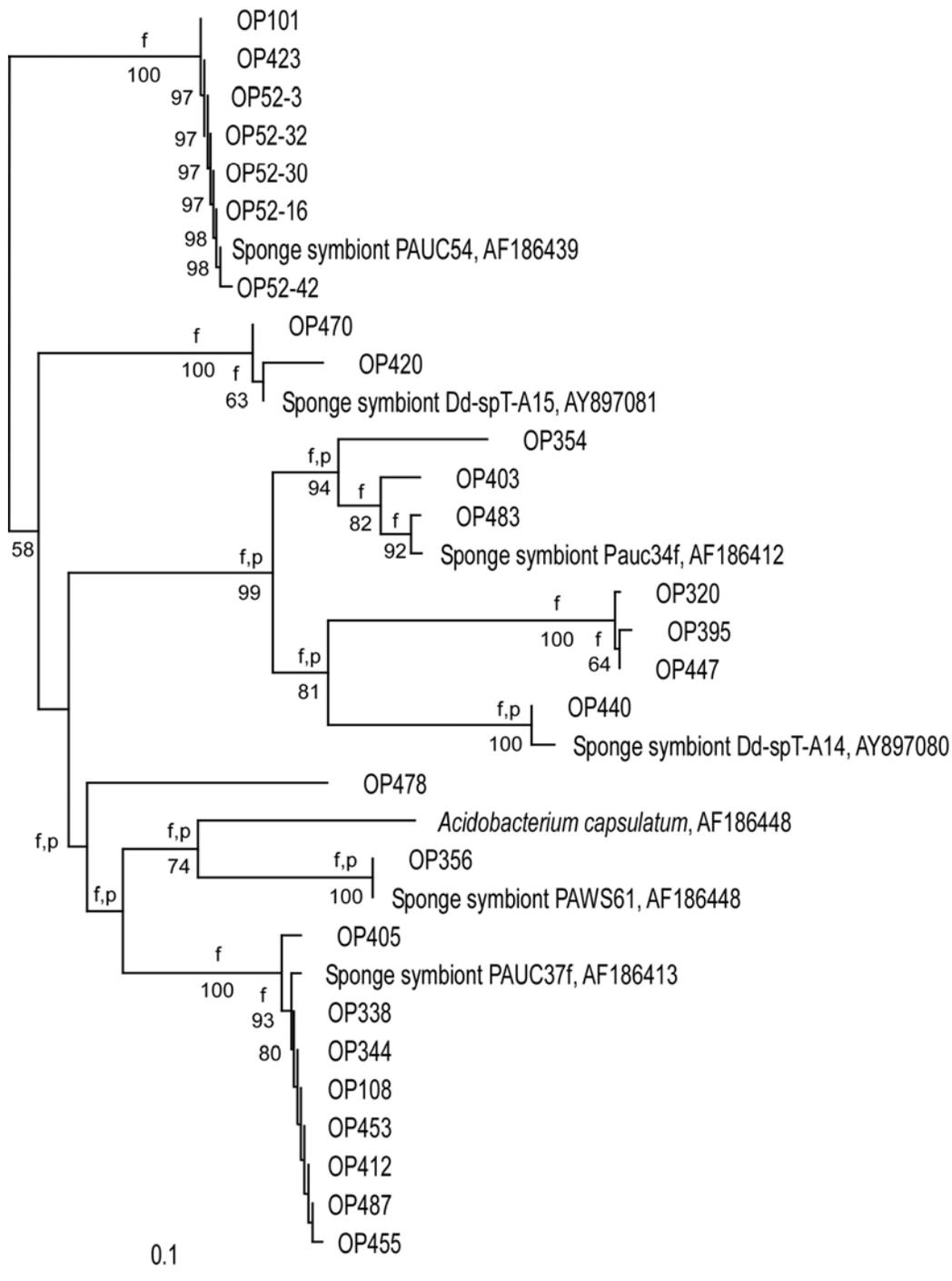
Figure 2.10. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with the *Acidobacteria* (Subgroup IV). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position. *E. coli* was used as the outgroup.
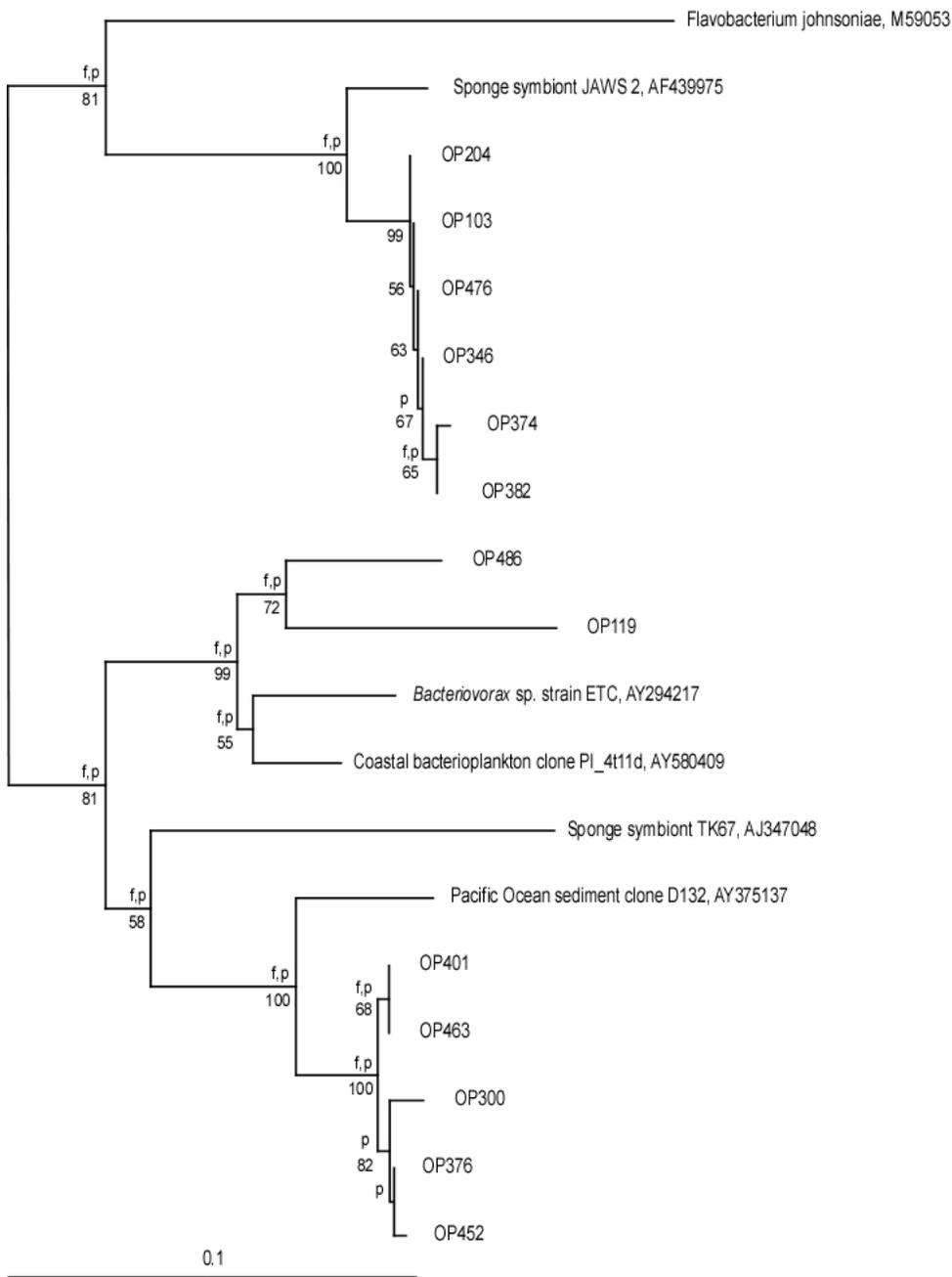
Figure 2.11. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with the delta-proteobacteria and *Bacteroidetes* (Subgroup V). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position. *E. coli* was used as the outgroup.
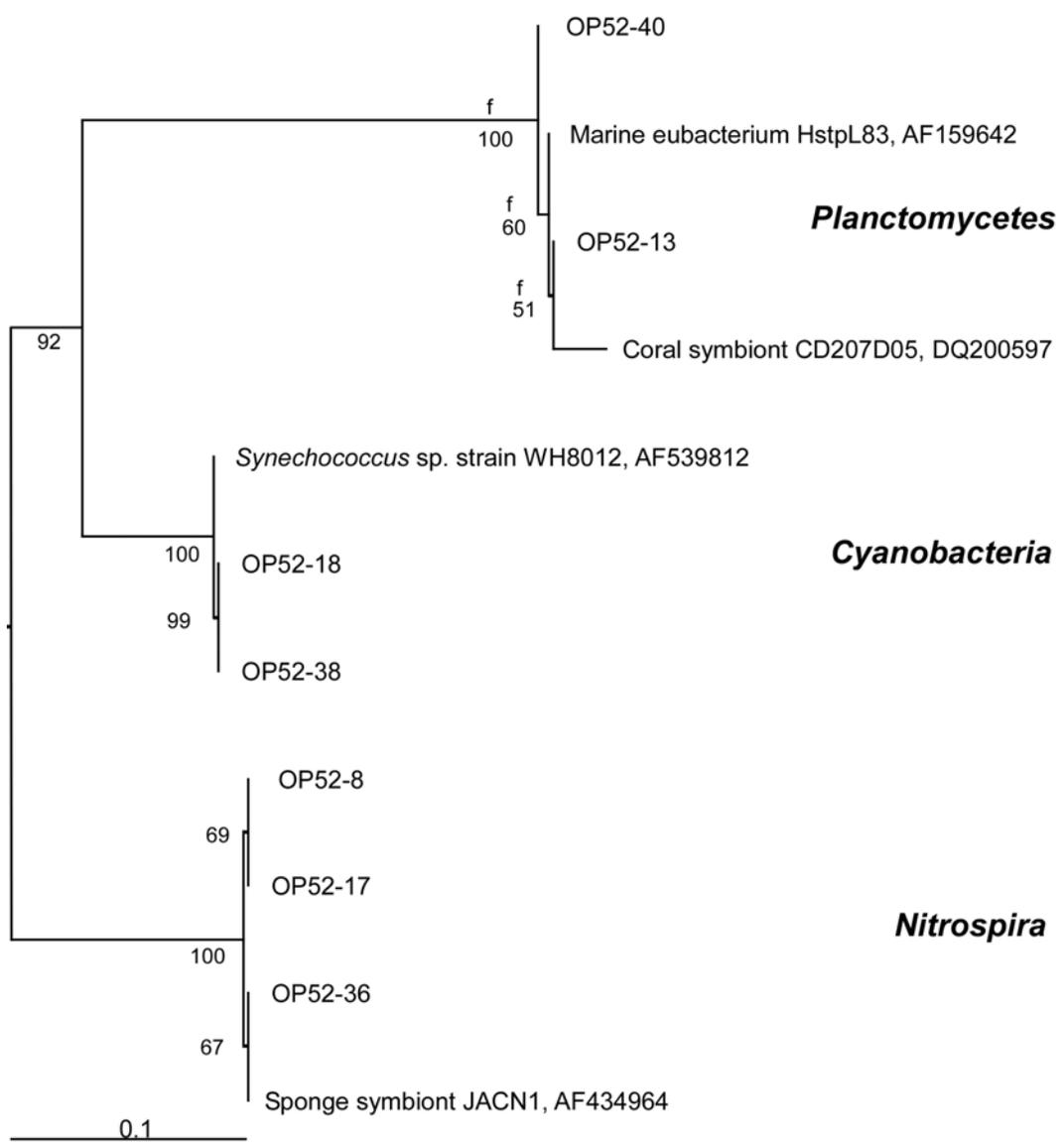
Figure 2.12. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from clones clustering with *Chloroflexi* and *Planctinomycetes* and *Nitrospira* (Subgroup VI). f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position. *E. coli* was used as the outgroup.

2.3.1.2 Sponge *Acanthostrongylophora* sp. Sponge 35 archaeal diversity indicated by 16S rRNA gene clone library analysis

PCR performed on the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 with primers specific for Archaea yielded a product of the expected size of 930 bp. The PCR product was cloned using TOPO-XL vector (Invitrogen) and yielded 60 colonies of which 11 were picked and sequenced. The sequence analysis revealed that two phylotypes, OPA3 and OPA11, were present and that they are crenarchaeotes (Figure 2.9). The two phylotypes clustered together and are distantly related to archaeal symbionts Ar23-23 (AY192641) and Ar19-36 (AY192640) from Korean sponges (Lee *et al.*, 2003).
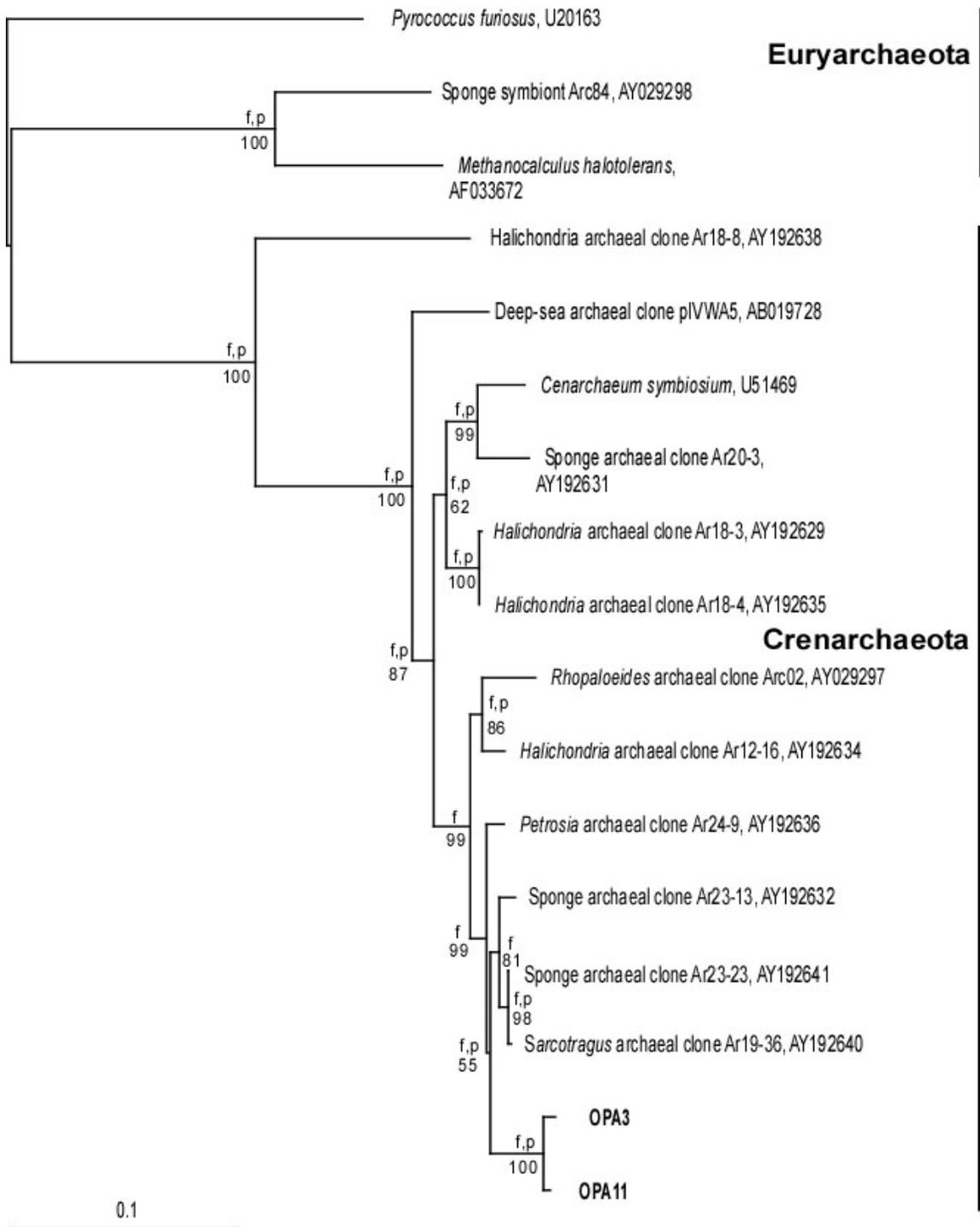
Figure 2.13. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from archaeal clones. f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Scale bar represents 0.1 substitutions per nucleotide position.

2.3.1.3 Fluorescent *in situ* hybridization studies

Using different group specific labeled probes, we were able to confirm and visualize the presence of some of the bacterial groups present in the 16S rRNA gene clone library. The EUB 338 probe, specific to most bacteria, showed an abundance of bacterial cell present in the mesohyl of *Acanthostrongylophora* sp. Sponge 35 (Figure 2.10A). The probe GAM 42a confirmed the presence of gamma-proteobacteria within the sponge tissue (Figure 2.10B). The probe HGC 69a, specific to high GC content bacteria, highlighted the presence of *Actinobacteria* in the mesohyl region of sponge (Figure 2.10C). Under the microscope, *Actinobacteria* formed clear clusters of about 20 to 50 bacteria not clearly distributed in specific part of the sponge tissue, but rather present in patches throughout the mesohyl region.
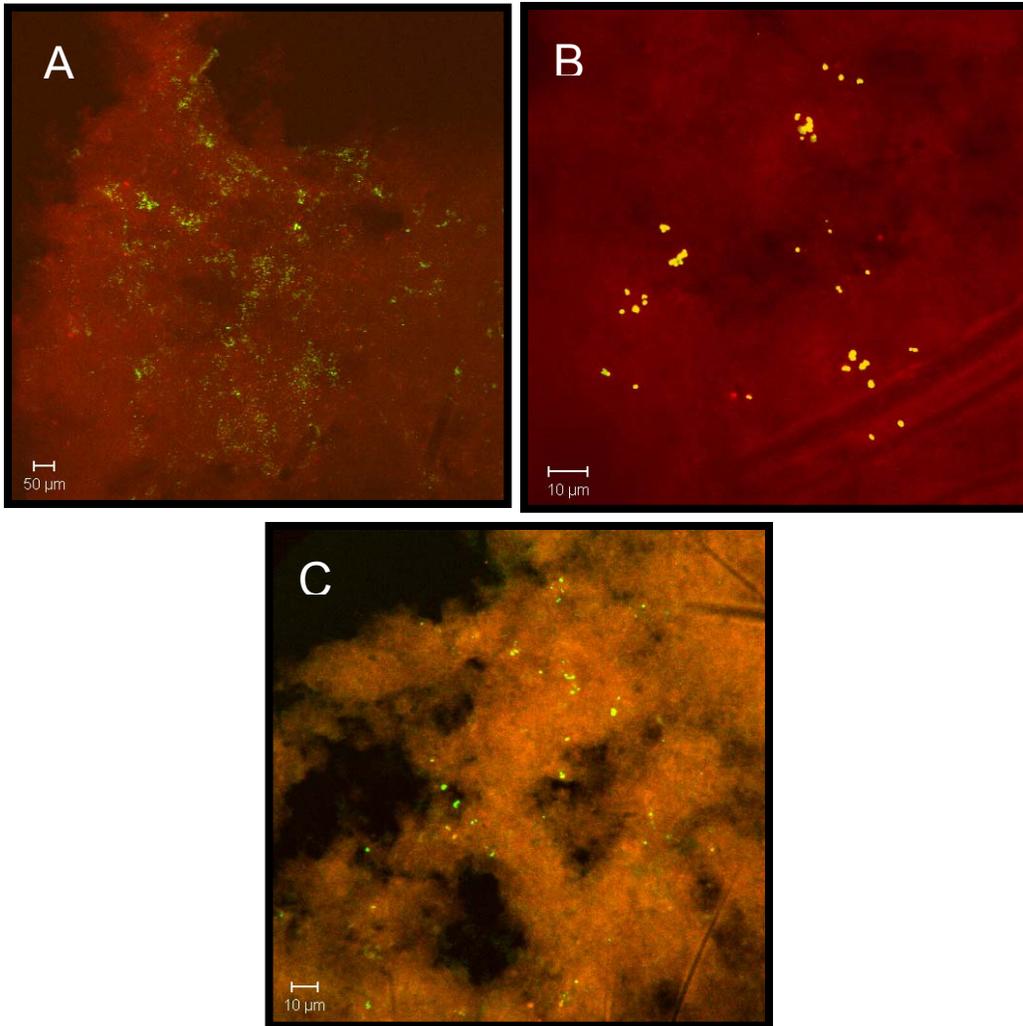
Figure 2.14. Epifluorescence micrograph section of *Acanthostrongylophora* sp. Sponge 35 mesohyl tissue visualized by FISH. The mesohyl region was hybridized with Cy-3 labeled EUB338 probe (A), Cy-3 labeled HGC69a probe (B) and Cy-3 GAM42a labeled probe (C).

2.3.2 DGGE analysis of the Indonesian sponges

Both sponge samples showed a band pattern clearly different from that of the surrounding water sample (Figure 2.14). The banding pattern for bacteria from *Acanthostrongylophora* sp. Sponge 35 is more complex than that of the seawater sample. The *Acanthostrongylophora* sp. Sponge 52 profile, though very different from that of the water, appears to be less complex. Comparing the two sponges, we can observe that they share several bands, two being especially noticeable at the bottom of the gel where high GC organisms would be located.
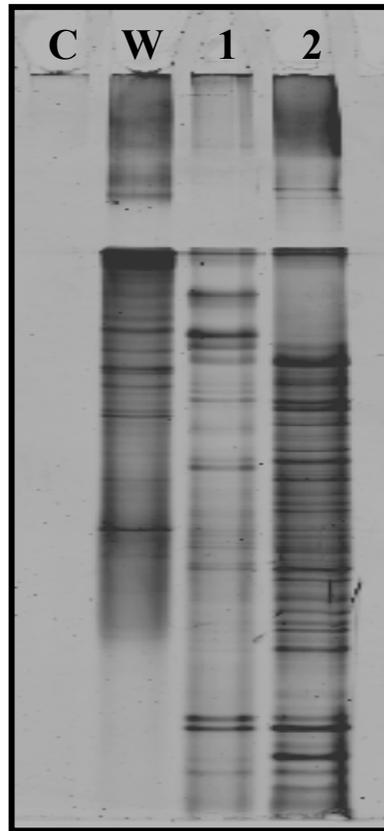
Figure 2.15 DGGE analysis of seawater (W), *Acanthostrongylophora* sp. Sponge 52 (1) and *Acanthostrongylophora* sp. Sponge 35 (2). C represents the PCR control.

2.4 Discussion

The analysis of two Indonesian sponge 16S rRNA gene clone libraries showed that both microbial communities were complex and diverse. Major bacterial groups shared by both sponges include *Actinobacteria, Acidobacteria, Chloroflexi*, alpha-proteobacteria and gamma-proteobacteria.

The microbial community of *Acanthostrongylophora* sp. Sponge 35 comprises eight different bacterial groups and is dominated by *Acidobacteria* and *Actinobacteria*. A rarefaction analysis of *Acanthostrongylophora* sp. Sponge 35 clone library showed that it was approprietly sampled at the class level. At the species level, the rarefaction analysis suggests that a greater diversity could be achieved by sampling additional clones.

The microbial community of *Acanthostrongylophora* sp. Sponge 52 comprises eight different bacterial groups. Some of the groups, *Nitrospira, Planctomycetes* and *Cyanobacteria* were not represented in the microbial community of the sponge *Acanthostrongylophora* sp. Sponge 35. The rarefaction analyses indicated that *Acanthostrongylophora* sp. Sponge 52 clone library was undersampled. At the class level, it would be likely that two other bacterial groups would have been represented with twice as many clones sequenced. At the species level, the rarefaction analysis suggests that the sequencing of 20 additional clones could result in ca. ten new species. The microbial community of *Acanthostrongylophora* sp. Sponge 52 is dominated by gamma-proteobacteria with actinomycetes forming the second largest group.

The Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 and 52 clone libraries were divided in subgroups according to BLAST results of the clone sequences. Subgroup I contains sequences related to *Actinobacteria* and represents the second largest group of the library with 23% of total clones. The excellent track record of actinomycetes in producing bioactive compounds and the suspected microbial origin of manzamines (Kobayashi & Ishibashi, 1993) makes this group of particular interest when trying to identify the source of manzamines. A majority of the sequences, 29 out of 33, had sponge symbionts as their closest relative. The *Actinobacteria* group had representatives from the groups Actino I and III (Hentschel *et al.*, 2002) that are present in the other previously studied sponges *R. odorabile*, *T. swinhoei* and *Aplysina aerophoba*. During FISH observation of *Acanthostrongylophora* sp. Sponge 35 tissue, with the HGC69a probe, actinomycetes had a tendency to group together but it did not seem that a particular location within the mesohyl harbored most of the actinomycetes. An interesting cluster of the *Actinobacteria* group was formed by eight clones from Sponge 35 closely related to the *T. swinhoei* symbiont JAWS10 (AF434968). It is possible that the JAWS10 cluster constitutes a new actinomycete genus specific to sponges. Additionally, a large cluster was composed of thirteen clones, ten of which from Sponge 52, and were closely related to *Theonella* sponge symbiont TK99 (AJ347072). Subgroup II is composed of alpha- and gamma-proteobacteria and accounts for 27% of the total number of clones. Among the gamma-proteobacteria, one distinct cluster is exclusively composed of clones from Sponge 52 clone library, all closely related to *Stenotrophomonas maltophilia* (DQ077704), an infrequent human pathogen found in

aquatic environment, which is capable of degrading the herbicide thifensulfuron-methyl. It is unclear what the role of these bacteria is in the sponge, it is possible because of the ability of *Stenotrophomonas maltophilia* to degrade complex toxic molecules, that they carry a similar function and may be able to degrade manzamine A. In the sponge *R. odorabile*, FISH analysis revealed that gamma-proteobacteria were especially predominant in regions surrounding the choanocytes chambers and it was hypothesized that they may be involved in nutrient uptake (Webster *et al.*, 2001b). Gamma-proteobacteria in the Indonesian sponge *Acanthostrongylophora* sp. Sponge 35 were well represented throughout the mesohyl with no specific accumulation of bacteria around the choanocytes chambers. Alpha-proteobacteria are well represented within the clone library. Recently, evidence for acyl homoserine lactone (AHL) production by bacteria associated with sponges was reported (Taylor *et al.*, 2004). Several alpha-proteobacteria isolated from additional sponges were investigated for the production of AHL with success (N. Mohamed, personal communication). It is possible that in the sponge-associated microbial community, some members of the alpha-proteobacteria group are involved in cell to cell signaling.

Subgroup III is composed of *Chloroflexi* and *Spirochaetes*. *Chloroflexi* are known to be involved in nutrient removal and are capable of carbohydrate degradation (Bjornsson *et al.*, 2002). It is possible that they perform a similar role with the sponge helping in nutrient uptake and waste management. *Spirochaetes* are an interesting group to find represented in the sponge microbial communities because of a study on *Spirochaetes* symbionts from termite guts that revealed their ability to fix nitrogen (Breznak, 2002). In nitrogen-depleted environments such as coral reefs, the ability to acquire nitrogen would

provide an ecological advantage to sponges and *Spirochaetes* associated with the sponge are good candidates to fill that function.

Subgroup IV is composed of *Acidobacteria*. Little is known about *Acidobacteria* except that they are ubiquitous and are an important component of all sponge microbial communities that have been studied.

Subgroup V contains delta-proteobacteria and *Bacteroidetes*. Unlike the sponge *Aplysina cavernicola* (Friedrich *et al.*, 1999), delta-proteobacteria are not the most abundant bacteria in the microbial community, representing 6% of the total number of clones. There have been reports of *Bacteroidetes* benefiting their host by degrading complex carbohydrates (Xu *et al.*, 2003) and also playing roles in the mineralization of organic matter and contributing significantly to waste water treatment (Okabe *et al.*, 2003). *Bacteroidetes* could help the sponge in the processing of metabolic waste.

The crenarchaeote sequences recovered had distant phylogenetic affiliation with three uncultured archaea from Korean sponges. The two crenarchaeote sequences clustered closely together. Euryarchaea and crenarchaea have been reported in several sponge microbial communities, none to date has been cultured and their role within the microbial community remains to be identified. A limited diversity of archaea has been observed in the sponges studied, two for *R. odorabile* (Webster *et al.*, 2001a), three for *Sarcotragus* (Lee *et al.*, 2003), and one for *Axinella mexicana* (Preston *et al.*, 1996).

The DGGE analysis of *Acanthostrongylophora* sp. Sponge 35, *Acanthostrongylophora* sp. Sponge 52 and the seawater showed that the two sponges have a very different

banding pattern from that of the surrounding water and no common bands could be observed. The banding patterns of the two sponges are also very different from each other with *Acanthostrongylophora* sp. Sponge 35 exhibiting a much more complex profile than *Acanthostrongylophora* sp. Sponge 52 confirming the results obtained from the 16S rRNA clone libraries. Nonetheless, several bands appear to be common to the two sponges especially in the lower part of the gel where high GC organisms such as actinomycetes would be located.

Both *Acanthostrongylophora* sp. Sponge 35 and *Acanthostrongylophora* sp. Sponge 52 were studied because of the presence of manzamines in their tissue. The comparison of the two microbial communities showed that they share major bacterial groups including *Actinomycetes, Acidobacteria, Chloroflexi,* alpha- and gamma-proteobacteria. The differences observed at the group level appear to be characteristic of each sponge. Some very distinct clusters within the same subgroup were formed almost exclusively by clones from either Sponge 35 or Sponge 52 clone libraries, i.e. actinomycete group closely related to *Theonella* sponge symbiont TK99 or exclusively, i.e. gamma-proteobacteria group closely related to *Stenotrophomonas maltophilia*. The two *Acanthostrongylophora* sponges have fairly similar microbial communities at the class level but the phylogenetic analyses highlighted very clear differences in several subgroups. The limited sampling of Sponge 52 clone library makes it difficult to draw any conclusion as to what the extent and impact of these differences may be. A more in depth analysis of Sponge 52 clone library would likely bring answers as to how much difference there is between *Acanthostrongylophora* sp. Sponge 35 and 52 microbial communities. The DGGE

analysis of the two sponges brought a limited answer to that question and showed that more similarities that were seen in the 16S rRNA clone libraries especially for high GC organisms. Considering the similar chemical signature of these sponges, it is possible that the organisms common to both sponges, as shown by the DGGE analysis, may be responsible for the manzamines found in both sponges.

3 Culture of sponge-associated bacteria

3.1 Introduction

The association between sponges and microbes is extremely ancient (Wilkinson, 1983b). Sponge fossils from deposits in Guizhou Province, China, dated at almost 580 million years, showed bacteria-like micro-spheroids or subspheroids (Yin & Gao, 2000). It has been suggested that sponge-associated microbes benefit their host in various ways including direct incorporation of dissolved organic matter from the surrounding seawater (Wilkinson & Garrone, 1980), contribution to the translocation of photosynthate by symbiotic cyanobacteria (Wilkinson, 1983a), transportation of metabolites throughout the mesohyl and carbon fixation (Borowitzka *et al.*, 1988), stabilization of the sponge skeleton (Wilkinson *et al.*, 1981) and chemical defense (Unson *et al.*, 1994). Microbes have been tremendous contributors to the world of natural products starting with the discovery of penicillin produced by *Penicillium notatum* by Fleming in 1928. The relentless search for new drugs has brought microbiology to the front line to isolate bacteria capable of producing new bioactive compounds. Terrestrial microbes have been intensely studied and have provided very efficient and widely used drugs including tetracycline produced by *Streptomyces* spp. and pseudomonic acid produced by *Pseudomonas fluorescens* (Chain & Mellows, 1977); however the rediscovery of metabolites has become an issue.

The marine environment represents a vast and relatively untapped resource for new drugs (Hill, 2004). More bioactive compounds have been isolated from marine sponges than

any other marine invertebrates (Blunt *et al.*, 2003; Blunt *et al.*, 2004; Blunt *et al.*, 2005; Faulkner, 1986; Faulkner, 1987; Faulkner, 1988; Faulkner, 1990; Faulkner, 1991; Faulkner, 1993; Faulkner, 1994; Faulkner, 1996; Faulkner, 1998; Faulkner, 2000; Faulkner, 2001; Faulkner, 2002). Similarities between sponge bioactive compounds and metabolites known to be of microbial origin have lead to suggestions that, in numerous cases sponge-associated bacteria rather than the sponges themselves are responsible for the production of bioactive compounds (Hildebrand *et al.*, 2004; Jensen & Fenical, 1994; Proksch *et al.*, 2002). Promising drugs with pharmaceutical potential have been isolated from marine invertebrates (Munro *et al.*, 1999) including bryostatin (Pettit *et al.*, 1982) and discodermolide (ter Haar *et al.*, 1996). Unfortunately, there has been a supply issue in order to obtain enough of these metabolites to proceed even to the clinical trial stage. The use of microbiology provides a solution for the isolation of sponge-associated microbes. Not all sponge symbionts can be cultured but with information gathered from the molecular investigation of the microbial community, appropriate media can be used to isolate specific groups of bacteria especially well represented in the 16S rRNA gene clone library or bacterial groups with particularly good track records for production of bioactive compounds, such as actinomycetes. For sponges containing bioactive compounds suspected to be of microbial origin, it is important to put much effort into the microbiological investigation of the invertebrate. The rewards of the successful isolation of a microbial producer are numerous. It opens the way to a reliable and sustainable production of the bioactive metabolites and also permits the leap to scaleable production. It also provides an unrestricted access to the genomic information of the microbe, making it available for study and improvement. Improvements can include facilitating cloning of

the biosynthetic gene cluster and expression in heterologous hosts for enhanced

production or manipulation to produce new analogues of the compound of interest.

3.2   Material and Methods

3.2.1   Sampling

Specimens of the sponges *Acanthostrongylophora* sp. Sponge 35 and Sponge 52

were collected by SCUBA diving from reef slopes and vertical surfaces of Black Reef

Point (Manado Bay, Indonesia) at depth between 6 and 33 m. Sponges were transferred

directly to plastic bags containing seawater to prevent contact of sponge tissue with air.

In the laboratory, sponge tissue was kept at -80˚C until used. The sponge tissue was then

freeze-dried prior to DNA extraction.

3.2.2   Isolation of sponge-associated bacteria

Sponge extract was obtained by grinding 1 cm^3 of sponge tissue from

Acanthostrongylophora Sponge 35 in sterile artificial seawater using a mortar and pestle.

The extract was plated on three different media types targeting different bacterial groups.

Marine agar 2216 (Difco, Detroit, Mich) is a non-selective medium for growth of many

heterotrophic marine bacteria. ISP2 (Difco) and starch casein agar are media designed

for selective growth of *Actinobacteria*. These media were supplemented with a final

concentration of 10 μg of nalidixic acid ml^{-1}, 10 μg of cycloheximide ml^{-1}, 25 μg nystatin ml^{-1} and 2% NaCl. Nalidixic acid is active against Gram-negative bacteria and cycloheximide and nystatin are fungicidal. Serial dilutions of the extract were plated on the three media and cultures were incubated at 30ºC for seven days in the case of marine agar and up to 9 weeks for the actinomycete media. One morphotype of each bacterial colony was selected and subcultured. Additional actinomycetes were grown from frozen sponge material on ISP2, R2A, chitin agar, actinomycete isolation agar, sea water starch agar and marine agar 2216 media following the previously described method. Formulations for all media are given in Appendix 2.

3.2.3 DNA extraction of sponge-associated bacterial isolates

Bacterial isolates were grown at 30˚C, in 50 ml broth of the medium they were originally isolated from (i.e. marine broth 2216, ISP2 or starch casein). After 5 days of growth, cultures were harvested by centrifugation and DNA was extracted from the cell material using the Ultra-Clean Microbial DNA Isolation kit (Mo Bio Laboratories Inc., Carlsbad, CA). DNA was quantified by spectrophotometry before being used as substrate for PCR amplifications.

3.2.4 Phylogenetic analysis of sponge-associated bacterial isolates

Isolates were identified by 16S rRNA gene sequence analysis. The 16S rRNA genes from the isolates were amplified by PCR using primers 8-27f (Weisburg *et al.*, 1991) and 1492 r (Reysenbach *et al.*, 1992). Cycling conditions were a hot start at 94°C for 5 min, 20 cycles of 92°C for 30 sec, 48°C for 2 min, 72°C for 1.5 min, and a final extension step at 72°C for 5 min. Thermal cycling was performed in a PTC-200 cycling system (MJ Research, Waltham, Mass.). PCR products were purified by electrophoresis in a 1% (wt/vol) agarose gel and bands of approximately 1500 bp were excised and recovered using a gel extraction kit (Qiagen, Inc., Chatsworth, Calif.). Sequencing was done using an ABI 377 automated sequencer (PE Applied Biosystems, Foster City, Calif.) with the 8-27f forward primer. Sequences were compared to those in GenBank with the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) algorithm to identify known closely related sequences. Ten sequences, with a length of at least 500 bp, were manually compiled and aligned with Phydit software (Chun, 1995). Trees were generated by the neighbor-joining (Saitou & Nei, 1987) algorithm (Kluge & Farris, 1969) implemented in Phydit. The robustness of inferred tree topologies was evaluated after 1,000 bootstrap resamplings of the neighbor-joining data, and only values >50% were shown.

3.2.5 Chemistry analysis of sponge-associated bacterial isolates

Heterotrophic bacteria were grown in 2 l flasks, in 500 ml of their medium of
isolation. The flasks were shaken at 150 rpm and incubated at 30˚C for five days. After
five days, the cultures were harvested by centrifugation. The supernatants were freeze-
dried and sent along with frozen wet cell pellets to the laboratory of Mark T. Hamann
(University of Mississippi) to be tested for manzamines. Detection of manzamines was
conducted using LC-MS. LC-MS analysis was carried out on a Bruker microTOF with
electrospray ionization. Solutions of standard manzamine A from the sponge were
prepared in methanol and subjected to LC-MS using reverse-phase C8 column
(Phenomenex, 5μ, 4.6 X 15 mm) eluting at 0.4 ml/minute flow rate with 20 minute linear
gradient from 20% to 100% phase B. Phase A was water and phase B was acetonitrile
and contained 0.1% formic acid. The ^{1}H NMR spectra were recorded in $CDCl_3$ using
NMR spectrometers operating at 400 or 500 MHz.

3.3 Results

A total of 16 bacterial morphotypes were selected from the initial plating from a
total of 3.46 10^7 cells/ml, representing 0.45% of the total number of bacteria estimated by
DAPI counting. After subculturing, ten pure isolates were recovered. Marine agar 2216
yielded eight isolates (M28, M29, M31, M34, M36, M37, M39, M40), starch casein one
isolate (M41) and ISP2 one isolate (M42). Phylogenetic analysis showed that isolate
M37 is a gamma-proteobacterium closely related to *Microbulbifer cystodytense*

(AJ620879) (Figure 3.1). Isolates M31 and M36 are alpha-proteobacteria closely related

to *Pseudovibrio denitrificans* strain DN34 (AY486423) and *Pseudovibrio* sp. strain

MKT94 (AB112827) respectively (Figure 3.1). Isolates M28, M29, M34, M39, M40 are

firmicutes (Figure 3.2). Finally, isolates M42 and M41 are actinomycetes. M42 is

closely related to *Micromonospora* sp. strain e17 (AY360157) and M41 is closely related

to *Microbacterium barkeri* (X77446) (Figure 3.3). One additional actinobacterium was

grown from frozen sponge material on R2A medium and named M62. After

phylogenetic analysis, M62 was identified as a *Streptomyces* sp. closely related to

Streptomyces violaceusniger (AJ391823) (Figure 3.3).
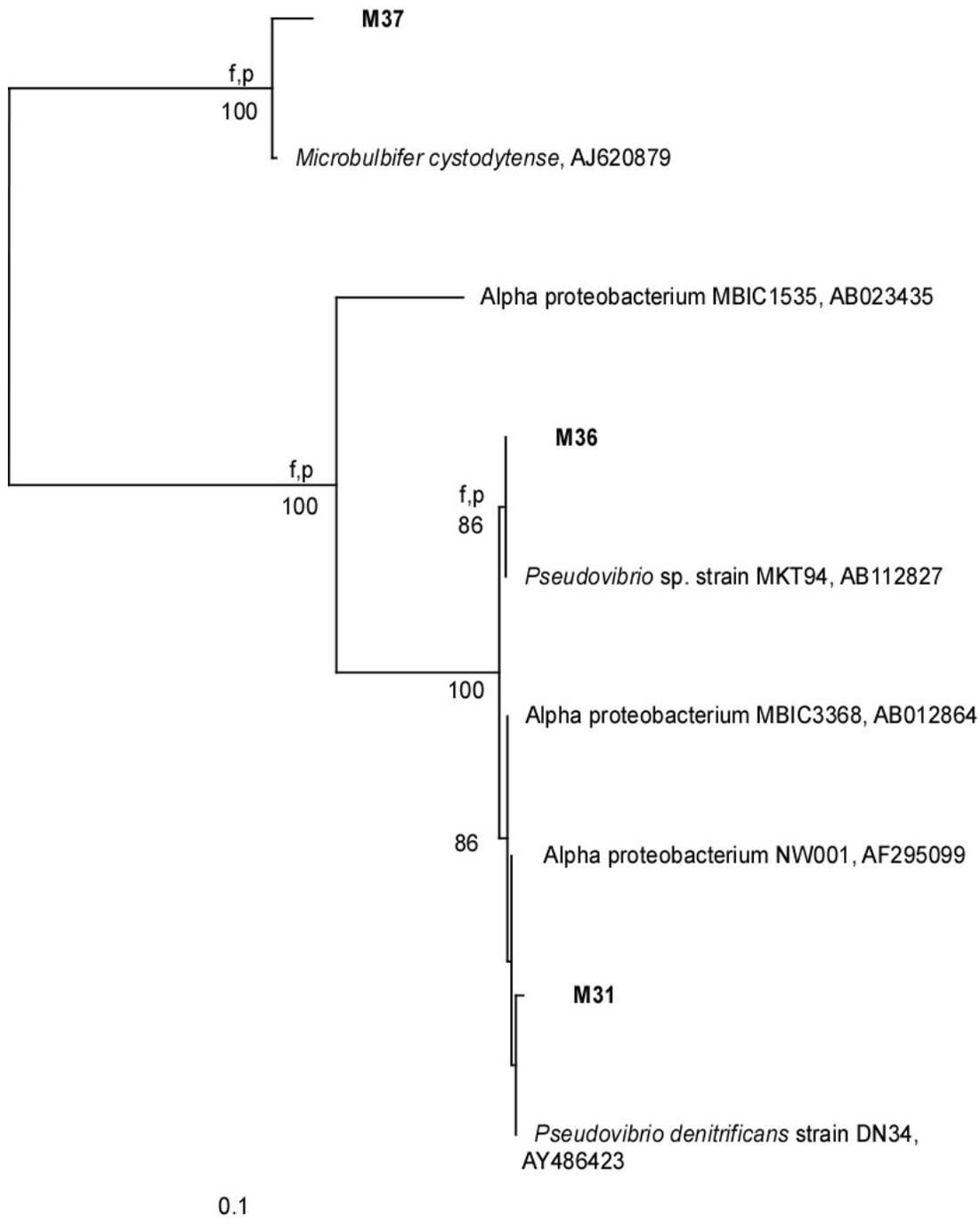
Figure 3.1. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from heterotrophic alpha-proteobacteria isolated from *Acanthostrongylophora* sp. Sponge 35. f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Only values > 50% are shown. Scale bar represents 0.1 substitutions per nucleotide position.
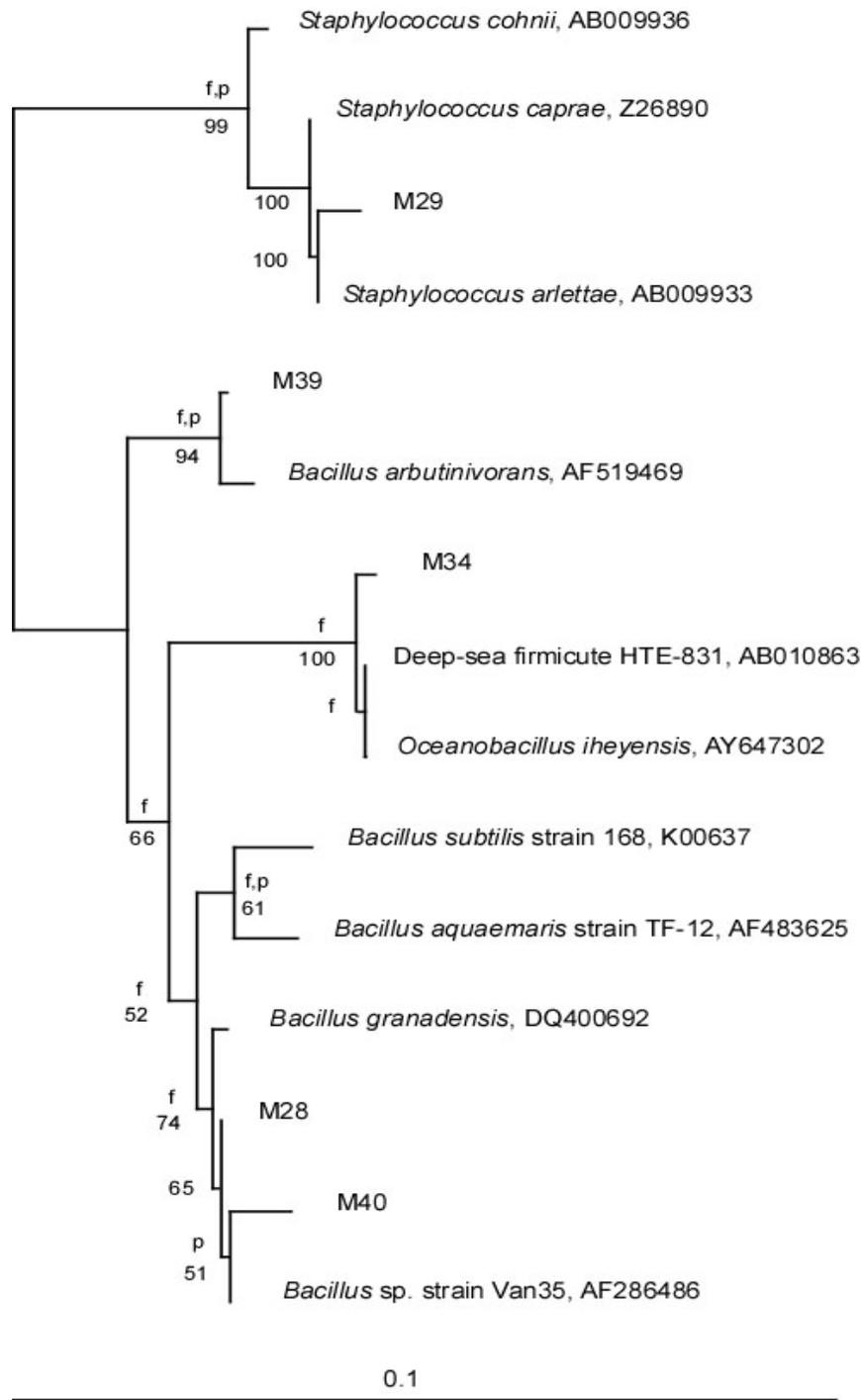
Figure 3.2. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from heterotrophic firmicutes isolated from *Acanthostrongylophora* sp. Sponge 35. f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Only values > 50% are shown. Scale bar represents 0.1 substitutions per nucleotide position.
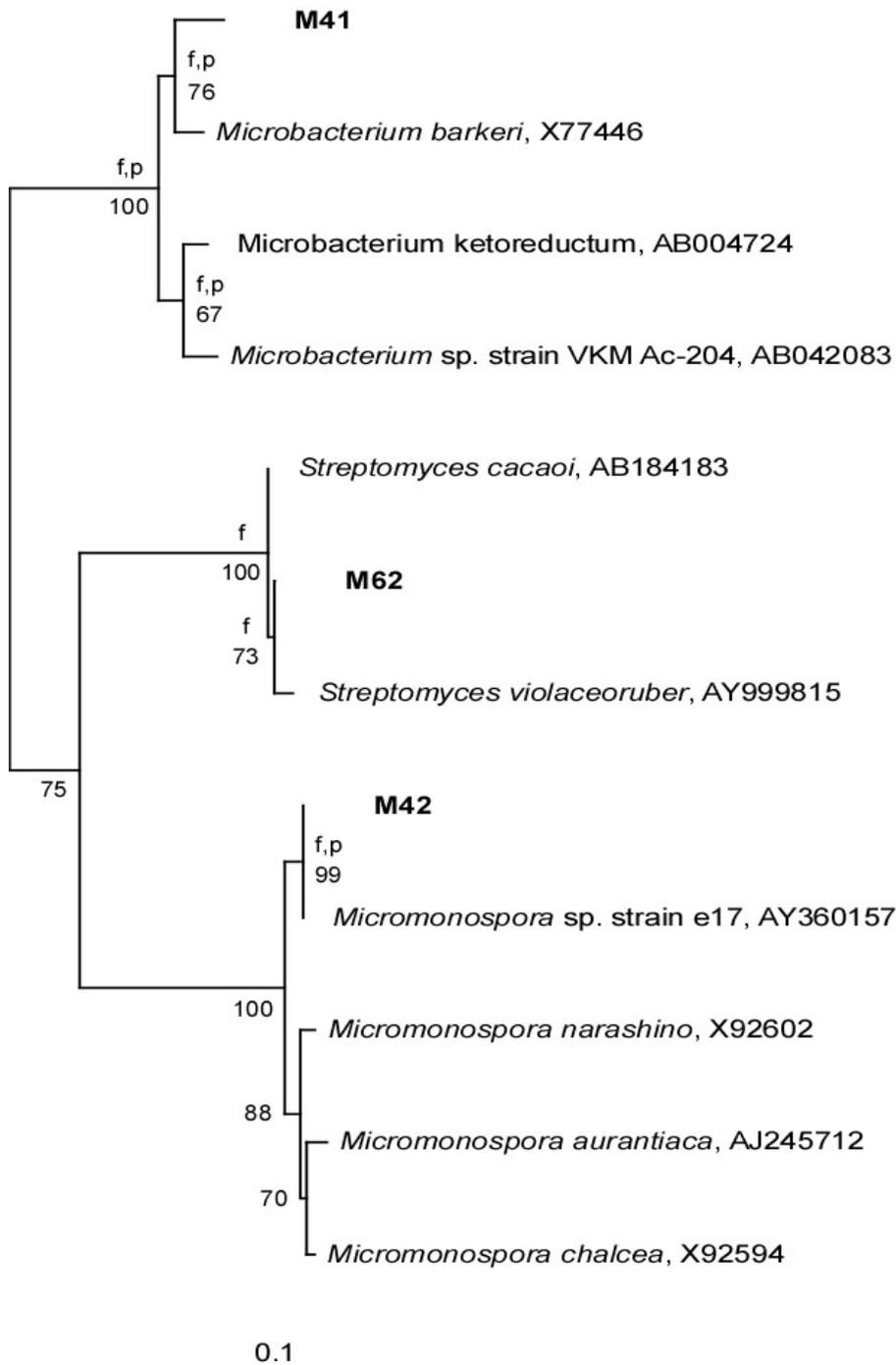
Figure 3.3. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence from heterotrophic *Actinobacteria* isolated from *Acanthostrongylophora* sp. Sponge 35. f and p indicate branches that were also found using the Fitch-Margoliash or maximum parsimony methods, respectively. The numbers at the nodes are percentages indicating the levels of bootstrap support, based on a neighbor-joining analysis of 1,000 re-sampled data sets. Only values > 50% are shown. Scale bar represents 0.1 substitutions per nucleotide position.

Each heterotrophic bacterium was tested for the presence of manzamines and compared to manzamine A (Figure 3.4 A) and 8-hydroxy-manzamine (Figure 3.4 B) standards obtained from the *Acanthostrongylophora* sp. Sponge 35 using LC-MS. Results from the LC-MS analyses of the different isolates cultures revealed the presence of manzamines in *Micromonopora* sp. strain M42. Manzamine A (Figure 3.4 C) and 8-hydroxy-manzamine isolated from *Micromonospora* sp. strain M42 are comparable to manzamine standards isolated from the sponge.
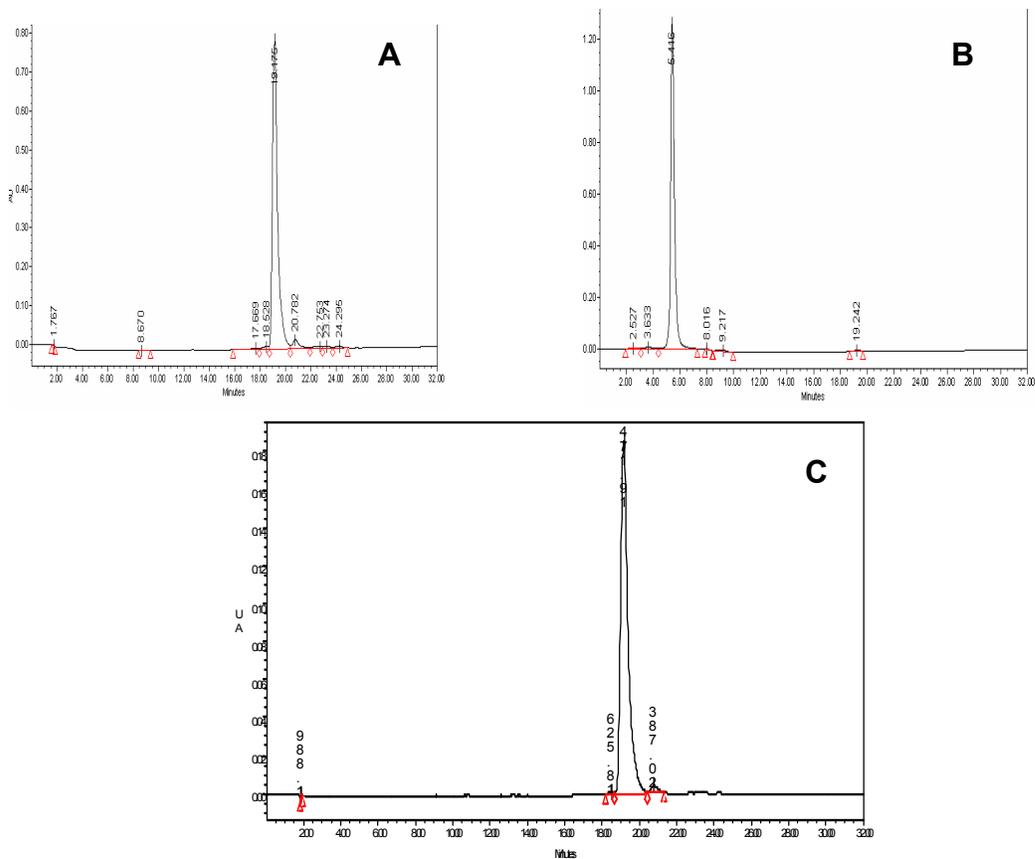
Figure 3.4. LC-MS chromatographs of manzamine A (**A**) and 8-OH manzamine A (**B**) obtained from *Acanthostrongylophora* sp. Sponge 35 and manzamine A obtained from *Micromonospora* sp. strain M42 (**C**).

3.4 Discussion

The number of culturable bacterial cells grown from the sponge fell within the expected range (<1%) (Buck, 1974) and major bacterial groups found in other sponges, including alpha- and gamma-proteobacteria, firmicutes and actinomycetes are well represented within these isolates. It is interesting to note that a variety of actinomycete genera were isolated from the sponge including *Micromonospora, Microbacterium* and *Streptomyces*. This result confirms the data obtained from the clone library that showed that a large part of the sponge microbial community consisted of diverse actinomycetes. With a more targeted effort, using a range of media specifically designed for actinomycetes, we were able to grow one additional actinobacterium. Work done on the Australian sponge *R. odorabile* (Webster *et al.*, 2001b) demonstrated that the use of wider variety of media aimed at a particular bacterial group increases the number of isolates that can be recovered when compared to the use of standard media. The authors used sponge extracts to supplement some media and observed that though it decreased the number of morphotypes, it stimulated the growth of a limited number of new morphotypes. Recently *Actinobacteria* belonging to the sub-class *Acidomicrobidae* were shown to be an important component of the sponge-associated microbial community of geographically distant *Xestospongia* spp. (Montalvo *et al.*, 2005). With specifically designed media containing various vitamins and trace elements, *Actinobacteria* from several genera were grown from *Xestospongia* spp. (N. Montalvo, personal communication). These examples demonstrate that it is possible to grow many more sponge-associated bacteria with appropriately designed media supplemented with

sponge-extracts, vitamins (Webster *et al.*, 2001b) or by simulating their natural environment (Kaeberlein *et al.*, 2002).

The alpha-proteobacteria grown from the sponge were of interest because the detailed study previously conducted on *R. odorabile* that showed the culturable bacterial community was dominated by the alpha-proteobacterium NW001 (Webster & Hill, 2001). On the marine agar 2216 medium, M31 had a morphotype similar to that of the alpha-proteobacterium NW001. The phylogenetic analysis showed that M31 was closely related to alpha-proteobacterium strain NW001. Bacteria closely related to this strain have been recovered from seven different sponge species geographically distant from one another (Enticknap *et al.*, 2006). Furthermore, it was demonstrated that a closely related strain to NW001 isolated from *M. laxissima* was transfered vertically through the larval masses, providing insight as to how sponges can acquire bacterial symbionts (Enticknap *et al.*, 2006). Though no role has been ascribed to strain NW001, a study on a diseased *R. odorabile* showed that strain NW001 was absent from this sponge culturable microbial community (Webster *et al.*, 2002), whereas it dominates the sponge culturable microbial community of healthy sponges. Strain NW001 could then be proposed as a bio-indicator of sponge health in *R. odorabile*.

With the study of more and more sponge-associated microbial communities, a trend has emerged concerning the phylogenetically close relationship between sponge-associated bacteria of different and geographically distant sponge species. The first observation of this phenomenon was made when comparing the 16S rRNA gene clone library of five different sponges (Hentschel *et al.*, 2002). This comparative study found that sponge-associated bacterial communities were phylogenetically diverse but also that there was a

sponge-specific microbial community with similar bacterial clades common to different sponge species and to single species found in different oceans. Recently, Enticknap *et al*. (2006) showed that this trend could be extended to sponge bacterial isolates often not represented in clone libraries or water column samples.

In the example of strain NW001, this alpha-proteobacterium was not represented in the clone library nor were any other alpha-proteobacteria present (Webster *et al.*, 2001b). Yet, fluorescent in-situ hybridization using a specific NW001 labeled probe showed that it was the predominant and possibly the sole alpha-proteobacterium, present in the sponge *R. odorabile* (Webster & Hill, 2001). The cell density of strain NW001 within the sponge was estimated to be between 10^4 and 10^6 cell/ml (Webster & Hill, 2001). The discrepancy between the presence on plates of culturable microbes and their absence in the 16S rRNA clone library can be due to multiple factors. The abundance of the microbe in the sample may be too low to be detected. The DNA extraction with its lysis stage can have a different effect on the bacteria, some being resistant to the lysis and may end up being underrepresented, others may be overly sensitive to it resulting in degraded DNA that is difficult to amplify. The number of 16S rRNA genes within each organism also plays an important role as the next step is a PCR amplification that will create a bias towards organisms with more 16S rRNA gene copies. The type of primers used for the PCR amplification of 16S rRNA genes creates another bias when trying to assess an entire microbial community as the so-called universal primers are not as universal as originally thought and can favor some species over others (Von Wintzingerode *et al.*, 1997).

Observations of sponge-associated microbial communities seem to indicate that sponges apply a strict selection in the acquisition of their symbionts. Though symbionts may vary between sponges, the fact that the same bacterial group and clusters are present within different sponge microbial communities could be interpreted less as a species-specific selection but rather as a task-specific selection of symbionts. It is possible that bacteria from a same cluster, capable of a similar symbiotic function, are interchangeable between different sponges.

The chemistry analyses indicated that *Micromonospora* sp. strain M42 produces manzamine A and 8-hydroxymanzamine which are two compounds previously detected in the sponge *Acanthostrongylophora* sp. The origin of manzamines has been proposed to be microbial (Kobayashi & Ishibashi, 1993) and the isolation of *Micromonospora* sp. strain M42 unequivocally confirms this hypothesis. The isolation of this manzamine producing strain opens new possibilities as to the production of a cost-efficient anti-malaria drug.

Many efforts have been devoted in isolating invertebrate-associated microbes producing the metabolites originally alleged to their host. A *Micrococcus* sp. bacterium isolated from the sponge *Tedania ignis* produced three diketopiperazines that were previously isolated from the sponge (Stierle *et al.*, 1988). Yet, it was argued that diketopiperazines are very common protein degradation products making it difficult to determine if these were primary compounds or just degradation products. A *Vibrio* sp. isolated from the sponge *Hyatella* sp. produces andrimid, an antibacterial mixed polyketide-peptide that was also detected in the sponge (Oclarit *et al.*, 1994), making it the first report of a

sponge-associated bacterium that produces a compound previously found in the host sponge.

Our isolation of *Micromonospora* sp. strain M42 confirms that bacterial symbionts can be the source of the bioactive compounds found in their host sponge and is the first report of a cultured symbiont producing a compound with important biomedical potential. This result demonstrates that microbiological approaches can provide a reliable solution to the supply issue often encountered with marine natural products.

4 Study of the manzamine-producing sponge-associated bacterium *Micromonospora* sp. strain M42

4.1 Introduction

The origin of metabolites in invertebrates has been the focus of much interest because of their activities and potential applications. The similarity of some bioactive compounds isolated from invertebrates with those from bacteria has led to speculation in many cases of a microbial origin (Hildebrand *et al.*, 2004; Jensen & Fenical, 1994; Proksch *et al.*, 2002). The sponge-associated bacterium *Micromonospora* sp. strain M42 tested positive for the production of manzamine A and 8-hydroxymanzamine. The isolation of the manzamine-producer *Micromonospora* sp. strain M42 is the second report of the cultivation of marine invertebrate symbiont that produces a bioactive compound found in the host and the first where the compound has great biomedical potential. Until this discovery, evidence of the involvement of symbionts in the production of their host metabolite was mainly the result of indirect methods or molecular investigations. Indirect methods such as cell sorting and chemical tagging have been used to co-localize bioactive compounds with bacteria (Unson & Faulkner, 1993; Unson *et al.*, 1994). However, chemical tagging is limited to the identification of a family of compounds rather than a specific metabolite. Co-localization of a compound with a particular cell type is not conclusive proof that the compound is made by these cells since many secondary metabolites are excreted from the producing bacteria. It is

conceivable that a compound may be produced by sponge cells and then specifically accumulated and concentrated by bacterial cells or the reverse. Small molecule secondary metabolites function as hormones, chemical communication and defense and as a result must be transported from the cell into the environment to elicit the desired effect. The acsidian *Lisoclinum patella* contains the cyclic peptide patellamide that was suspected to possibly be of microbial origin. The ascidian harbors symbiotic cyanobacteria *Prochloron* spp. for which cultivation attempts are yet to be successful. Using a genomic approach, Schmidt *et al* (2005) constructed a cosmid library of *Prochloron dimdemni*, the symbiont of the acsidian *L. patella* and succesfully identified the genes responsible for patellamide biosynthesis. Working also on the identification of the patellamide gene cluster, Long *et al.* (2005) took a shotgun cloning approach that resulted in the heterologous expression of patellamide D in *E. coli*. In the case of the anti-cancer compound bryostatin I, strong evidence has linked the production of the bioactive compound to "*Candidatus* Endobugula sertula" a symbiont of the bryozoan *Bugula neritina* (Davidson *et al.*, 2001). However, "*Candidatus* Endobugula sertula" has not been cultivated and a genomic approach to identify the genes responsible for the biosynthesis of bryostatin remains to be completed. These three examples demonstrate how challenging it is to identify the microbe responsible for the production of the compound of interest and furthermore, how difficult the identification of the biosynthetic pathway is when the microbe can not be cultured.

In the case of manzamines, the isolation of *Micromonospora* sp. strain M42 provides a unique opportunity to gather knowledge about the strain, its requirements for growth and the production of manzamine A. Having the strain available also facilitates a genomic

approach. It is possible to size the genome of *Micromonospora* sp. strain M42 and explore this genome for additional natural products that it may contain. Finally, it offers the possibility to look for the cryptic manzamine pathway about which little is known by searching for putative enzymes involved in the biosynthesis of manzamine A using both molecular and biochemical techniques.

4.2 Materials and Methods

4.2.1 Microbiology

Growth monitoring methods. Two methods to monitor the growth of *Micromonospora* sp. strain M42 were compared. The first one consisted in measuring the optical density of cultures at 600 nm (OD_{600}) and the second relied on plate counts with dilutions of *Micromonospora* sp. strain M42 plated on ISP2 medium and incubated for 10 days at 30˚C. These methods were used in parallel with a time course experiment that was conducted for 16 days.

Time course experiment. The experiment was conducted for 16 days and time points were set every three days (days 1, 4, 7, 10, 13, 16). Four flasks of 2 l volume were prepared with 500 ml of ISP2 medium. Three flasks were inoculated with 10 ml of a dense culture of *Micromonospora* sp. strain M42 corresponding to an OD_{600} of 0.5 and a fourth flask was set as a control. At each time point, 50 ml of culture was taken and centrifuged. The supernatant was separated from the pellet and lyophilized using a

freeze-dryer (FTS Systems, Inc, Stone Ridge, NY). Both pellet and supernatant were

then analyzed by TLC, NMR and mass spectrometry to detect and quantify manzamines.

The growth of the organism was monitored by measuring OD_{600} and by plate counting.

Micromonospora **sp. strain M42 salt requirement.** An experiment was done to test the

effect of different salt concentrations on growth and manzamine production by

Micromonospora sp. strain M42. The experiment was conducted for six days with time

points every day. Five different ISP2 media with NaCl concentrations of 0%, 1%, 2%,

3.5% and 5% respectively (50 ml of medium in 250 ml flasks) were inoculated with 1 ml

of a dense culture of *Micromonospora* sp strain M42. The growth of the organism was

monitored by measuring OD_{600}. To measure manzamine A concentration, supernatant

was collected at day 1, 3 and 6 and extracted with ethyl acetate. Manzamine A

concentration was measured using LC-TOF-MS.

4.2.2 Detection of relatives of *Micromonospora* sp. strain M42 from

sponge total DNA

From an alignment of organisms closely related to *Micromonospora* sp. strain

M42, one primer named M42-16S (5'-GACCGTGAAAACCTGGGGC-3') was designed

to be specific for close relatives of *Micromonospora* sp. strain M42 and was used in

combination with universal eubacterial primer 1492r. First, a 16S rRNA gene

amplification with primers 8-27f and 1492r DNA amplification was carried out on total

DNA extracted from *Acanthostrongylophora* using Hi-Fi Platinum *Taq* (Invitrogen, San

Diego, Calif.). A nested PCR was then performed with primers M42-16S and 1492r as follow: initial denaturation at 94ºC for 5 min, 35 cycles of 94ºC for 1 min, 69ºC for 30 sec, 72ºC for 2 min, and a final extension of 5 min at 72ºC in a PTC-200 MJ-research thermal cycler. PCR products were purified by electrophoresis in a 1% (wt/vol) agarose gel and bands of the predicted amplicon size of approximately 1000 bp were excised and recovered using a gel extraction kit (Qiagen, Inc., Chatsworth, Calif.). Purified PCR products were cloned with a TOPO-XL cloning kit according to the manufacturer's instructions (Invitrogen).

4.2.3 Genome sizing

Restriction digest of *Micromonospora* sp. strain M42 genome.

Micromonospora sp. strain M42 DNA plugs for PFGE analysis were prepared as described elsewhere (Ravel *et al.*, 1998). Two restriction enzymes, *Vsp*I and *Ssp*I, were used. *Micromonospora* sp. strain M42 DNA plugs were rinsed in the appropriate restriction buffer for 6 to 12 h at 4ºC, transferred to 300 µl of ice-cold buffer containing 50 U of the restriction enzyme *Vsp*I or *Ssp*I and 50 mg of acetylated bovine serum albumin (Promega Co., Madison, Wis.) and kept at 4ºC for 2 h prior to incubation for 12 to 16 h at 37ºC.

PFGE. Plugs were subjected to PFGE by using a clamped homogenous electric field system (CHEF DR-III; Bio-Rad, Melville, N.Y.) in 0.5× TBE buffer (1× TBE is 98 mM Tris-HCl, 89 mM boric acid, and 62 mM EDTA) containing 100 µM of thiourea at 14°C

(Ray *et al.*, 1992). Two different running conditions were used, one to separate

fragments between 30 kb and 500 kb and one for fragments between 450 kb and 1.5 Mb.

The gel running conditions for fragments between 30 kb and 500 kb were a run time of

27 h, angle 120˚, gradient 6 V/cm, initial switch time 2.16 s, final switch time 44.69 s.

Running conditions for fragments between 450 kb and 1.5 Mb were a run time of 33 h 53

min, angle 120˚, gradient 6 V/cm, initial switch time 52.22 s, final switch time 147 s.

DNA was stained with ethidium bromide and digitalized with a FluorImager 573

(MolecularDynamics, Sunnyvale, Calif.) and molecular weights were estimated with

LabImage Version 2.7 (Kapelan Bio-Imaging Solutions).

4.2.4 Manzamine biosynthetic pathway

PCR amplification of tryptophan decarboxylase. To amplify the putative tryptophan

decarboxylase gene, six PCR primers, two forward and four reverse primers, were

designed from an alignment of known tryptophan decarboxylase and aromatic amino acid

gene sequences. The sequence of the primers was matched to the codon usage of

Streptomyces and *Micromonospora* (Table 4.1). For the PCR, a mix of forward primers

TDC10 and TDC11 (50/50) and reverse primers TDC12 and 13 or TDC14 and 15

respectively, were used. The DNA amplification was carried out using Hi-Fi Platinum

Taq (Invitrogen, San Diego, Calif.). Cycling conditions were as follow: initial

denaturation at 94ºC for 5 min, 30 cycles of 94ºC for 1 min, 42ºC for 30 sec, 72ºC for 2

min, and a final extension of 5 min at 72ºC in a PTC-200 MJ-research thermal cycler.

PCR products were purified by electrophoresis in a 1% (wt/vol) agarose gel and bands

were excised and recovered using a gel extraction kit (Qiagen, Inc., Chatsworth, Calif.).

Sequences were then compared to those in the GenBank database using BLAST (Altschul

et al., 1990).

Purification of a putative tryptophan decarboxylase enzyme. For total protein

extraction, 5 g of *Micromonospora* sp. strain M42 cell pellet obtained from a culture

grown in ISP2 medium was used. The extraction was carried out using the Cell Lytic B

Plus kit (Sigma). The purification of the tryptophan decarboxylase enzyme was done

following a modified method by Nakazawa *et al.* (1981) (Appendix 3). A tryptophan

decarboxylase assay was set up to monitor enzyme activity as follows: 250 µl of 0.05 M

potassium phosphate, 20 µl of 5 10^{-3} M pyridoxal phosphate, 40 µl of tryptophan and 250

µl of M42 protein extract were incubated at 30˚C and samples were taken after 1 h, 2 h, 3

h and overnight. The samples were then run on a high pressure liquid chromatography

system (HPLC) (Agilent) and compared with tryptophan and tryptamine standards. The

HPLC mobile phase was water/acetic acid 0.01% and acetonitrile with a gradient starting

at 20% acetonitrile.

4.2.5 Detection of polyketide genes

Degenerate PCR primers, MDPQQRf (5'-RTRGAYCCNCAGCAICG-3') and

HGTGTr (5'-VGTNCCNGTGCCRTG-3') (Kim *et al.*, 2005) were used to amplify the

beta-ketosynthase domain fragment within Type I polyketide synthase genes. A

touchdown PCR was performed as follow: 95˚C for 5 min, 11 cycles of 95˚C for 1 min,

60˚C to 40˚C for 1 min (-2˚C/cycle), extension at 72˚C for 1 min followed by 39 cycles of 95˚C for 1 min, 40˚C for 1 min, 72˚C for 1 min and final extension at 72˚C for 10 min. PCR products were purified by electrophoresis in a 1% (wt/vol) agarose gel and bands were excised and recovered using a gel extraction kit (Qiagen, Inc., Chatsworth, Calif.). Purified PCR products were cloned with a TOPO-XL cloning kit (Invitrogen) according to the manufacturer's instructions. Sequencing of clones was done using an ABI 377 automated sequencer (PE Applied Biosystems, Foster City, Calif.) with MDPQQRf primer. Sequences were compared to those in the GenBank database using the BLASTX algorithm to identify known closely related sequences. Additionally, one of the amplified PKS sequences was used as a probe for Southern hybridization of a pulsed field gel containing *Micromonospora* sp. strain M42 plugs digested with *Ssp*I, *Vsp*I, *Pac*I, *Dra*I and *Smi*I. The DNA was transferred to Managraph nylon membrane (MSI, Micron Separation Inc., Westborough, Mass.) following the recommended protocol of the manufacturer. The PKS probe was labeled with ^{32}P using All-In-One Nick translation DNA labeling mix (α-dCTP) (Sigma) according to the manufacturer's instructions. The hybridization was performed at 65˚C and washed at high stringency (68˚C) in 0.1X SSC, 1% SDS following the MSI protocol. Results were analyzed digitally using a Typhoon PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.).

4.2.6 Improvement of manzamine production by mutagenesis of

Micromonospora sp. strain M42

Micromonospora sp. strain M42 was grown in shaked flasks (120 rpm) and a 20 l

fermenter in ISP2 liquid medium with vigorous aeration at 30˚C. To generate mutants,

spores harvested after 10 days growth on ISP2 plates were irradiated using a 254 nm UV

source to achieve a 99% kill rate. Resultant colonies were inoculated into 50 ml ISP2

liquid cultures and shaken for 4 days at 30˚C. The samples of *Micromonospora* sp. strain

M42 were extracted with ethyl acetate at room temperature and extracts were screened by

thin layer chromatography (TLC) for improved manzamine production. Extracts of

mutants with improved manzamine production capabilities were then processed to

identify and quantify the compound as follow: the extracts were concentrated under

reduced pressure and the resultant concentrated extract was subjected to silica gel

vacuum-liquid chromatography and eluted beginning with hexane (100%), hexane–

acetone (9:1, 3:1, 1:1), acetone (100%), chloroform–methanol (1:1) and finally with

methanol (100%). Fraction 3, eluted with a gradient system of hexane–acetone (7:3) was

subjected to HPLC chromatography (Phenomenex Luna 5 μM 250 x 10.0 mm column,

flow rate 3 mL/min, λ 254, 360 nm) using a gradient solvent system of acetonitrile and

water to obtain manzamine A and manzamine B.

4.3 Results

4.3.1 Microbiology of *Micromonospora* sp. strain M42

The comparison of growth assessment methods for *Micromonospora* sp. strain M42 between optical density at 600 and plate counts showed that both methods generated fairly similar growth curves (Figure 4.1). The plate count method was able to give a precise count of colony forming units while the OD_{600} method was rapid and convenient and gave a good estimation of the organism's growth.
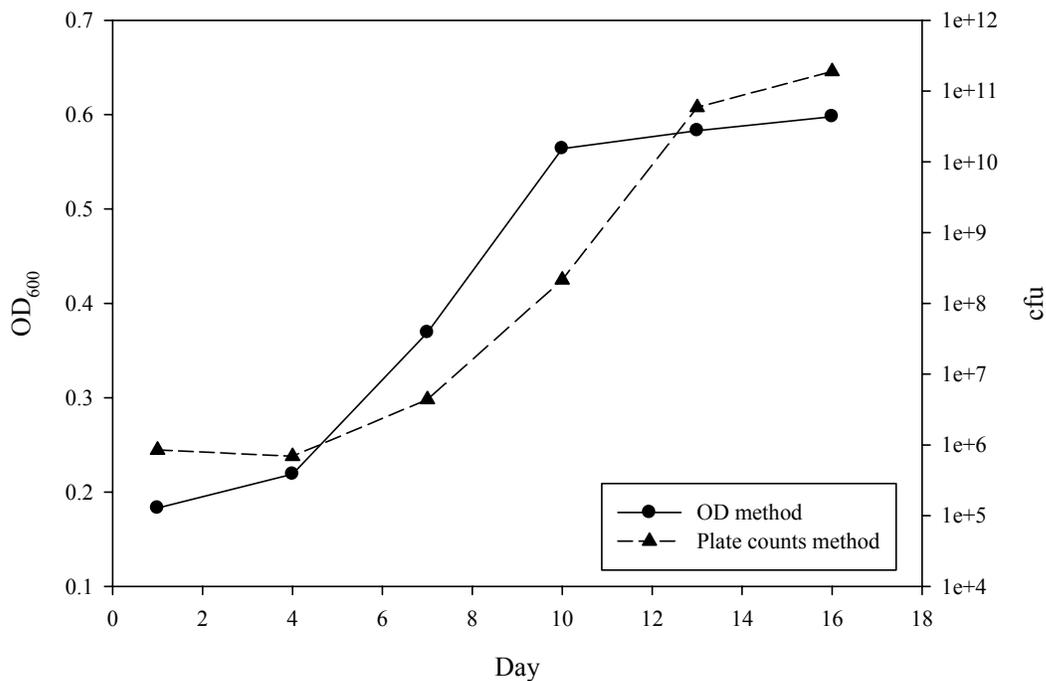
Figure. 4.1. Comparison between plate counts and OD_{600} methods to monitor the growth of *Micromonospora* sp. strain M42.

A 16 day time course experiment was set up to compare the growth pattern of

Micromonospora sp. strain M42 with the production of manzamine (Figure 4.2). The

results showed that manzamine is detected as early as one day after inoculation of the

medium with *Micromonospora* sp. strain M42. The manzamine concentration peaked

around day 7 which corresponds to the middle of the exponential growth phase of the

bacterium. At day 10, corresponding to the start of the plateau phase, no manzamine, or

manzamine degradation products, are detected in the medium.
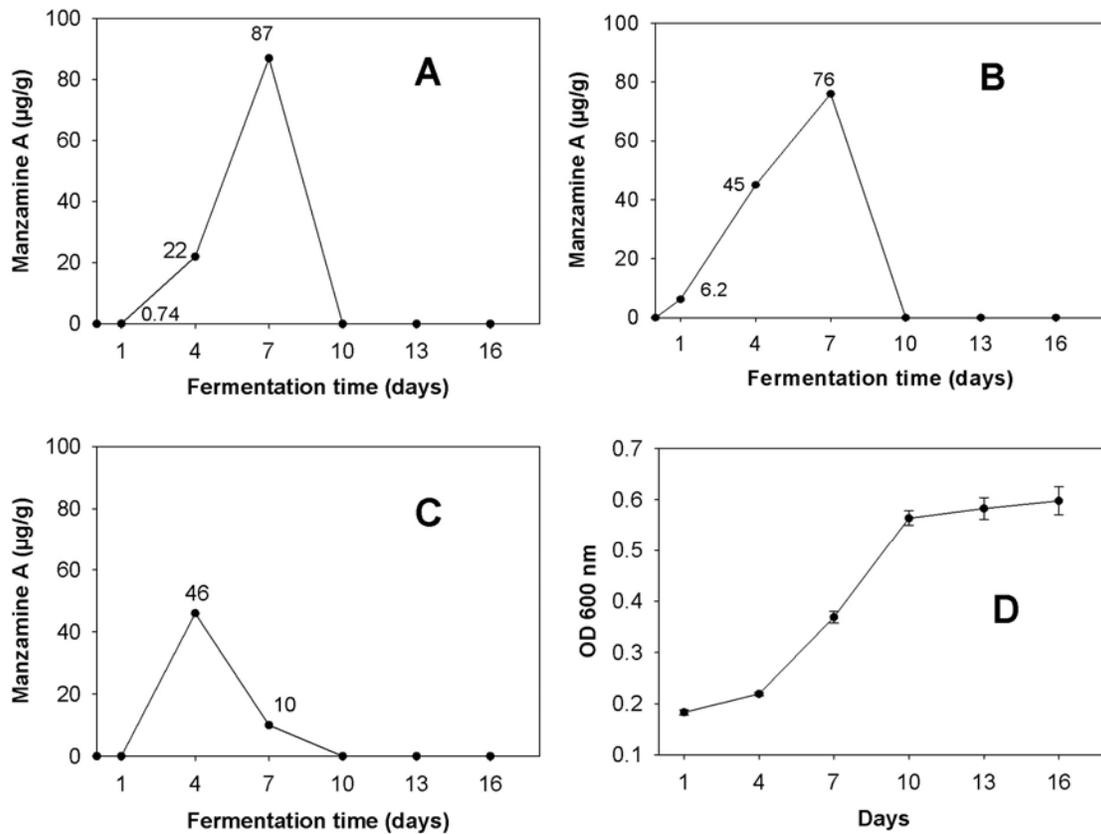
Figure 4.2. Quantification of manzamine A produced by *Micromonospora* sp. strain
M42 during a time course experiment. Panel A, B and C represent individual manzamine
A quantification data for three shaked flasks inoculated with *Micromonospora* sp. strain
M42. Panel D illustrates the growth of *Micromonospora* sp. strain M42.

Micromonospora sp. strain M42 was grown in ISP2 under five different salinity

conditions (0, 1, 2, 3.5 and 5 % NaCl). Results showed that *Micromonospora* sp. strain

M42 achieves higher cell density and faster growth at NaCl concentration between 0 and

2% (Figure 4.3). The highest cell density was reached with a NaCl concentration of 1%

and 0% provided the fastest initial cell growth. The three lower NaCl concentrations,

0%, 1% and 2% NaCl, yielded similar cell densities after 6 days. With a 3.5% NaCl

concentration, the growth of *Micromonospora* sp. strain M42 is delayed until day three

and rapid growth is not attained, resulting in a low cell density after six days. The

highest NaCl concentration tested, 5% NaCl, hampered the growth of the organism.

Manzamine A concentrations were monitored during this experiment. The results

showed that the manzamine A concentration matches the cell density of the cultures. The

highest manzamine A yield was obtained at 0% NaCl, followed by 1%, 2%, 3.5% and 5%

NaCl (Figure 4.4). In this experiment, the maximum concentration of manzamine A was

obtained at day 3 for all samples.
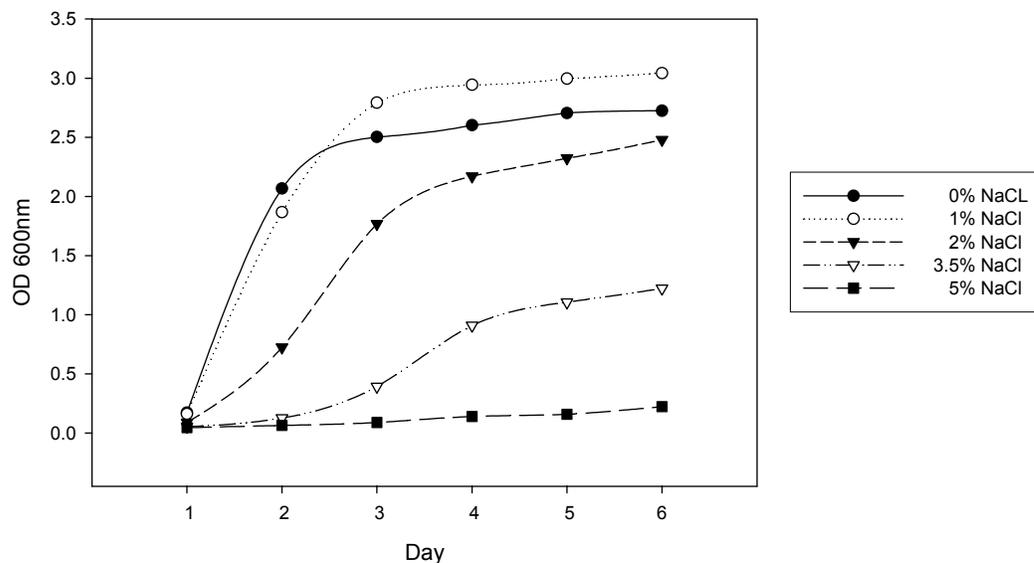
Figure 4.3. Effect of ISP2 medium containing different NaCl concentrations on growth
of *Micromonospora* sp. strain M42.
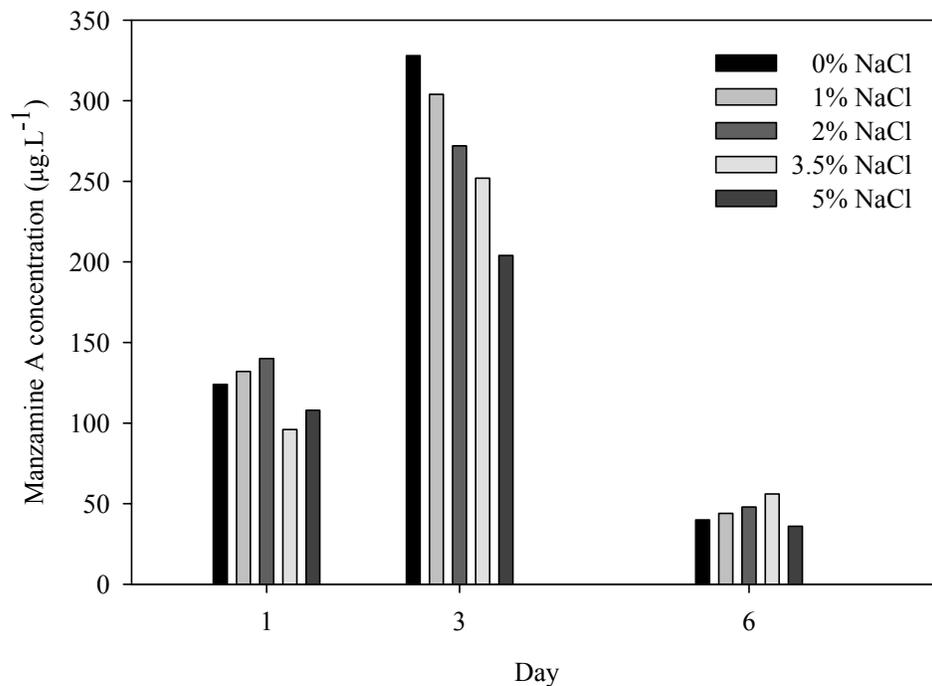
Figure 4.4. Production of manzamine A by *Micromonospora* sp. strain M42 in ISP2 medium containing different NaCl concentrations.
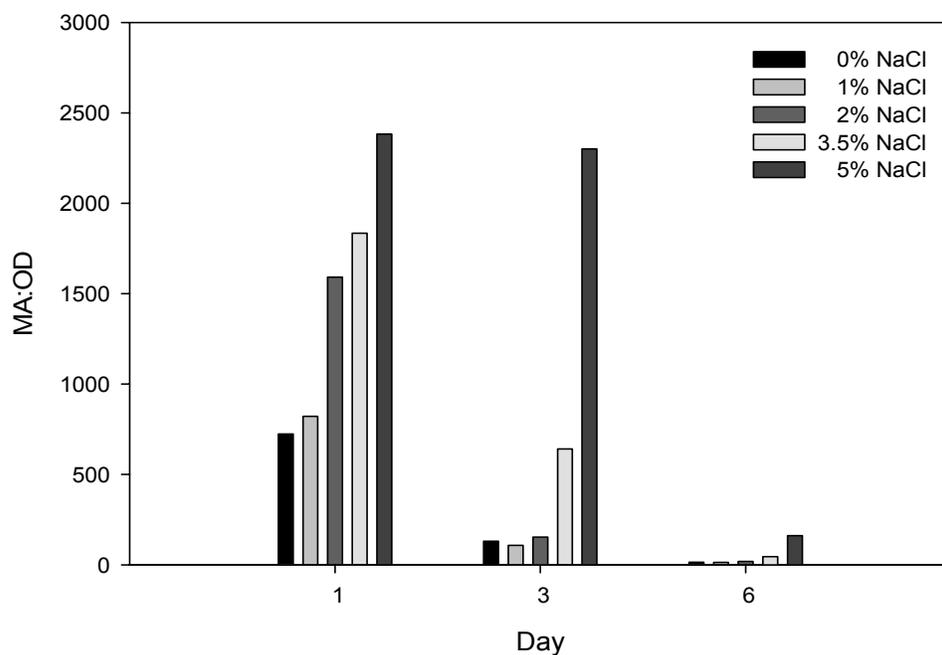
Figure 4.5. Comparison of the ratio of manzamine A to OD_{600} of *Micromonospora* sp. strain M42 cultures grown under different salt concentrations.

The ratio of manzamine A to OD_{600} was calculated to examine the relationship between

Micromonospora sp. strain M42 cell density and the amount of manzamine A produced

(Figure 4.5).

To detect the presence of strains closely related to *Micromonospora* sp. strain M42 in the

Acanthostrongylophora Sponge 35 extracted total sponge DNA, the designed primer

M42-16S and eubacterial primer 1492r were paired for PCR. No amplification was

visible when amplifying directly on total sponge DNA. A nested PCR approach was then

used and a PCR was performed using M42-16S/1492r primers on the 16S rRNA gene

PCR product amplified from *Acanthostrongylophora* Sponge 35 total sponge DNA using

the primer pair 8-27f and 1492r. This nested PCR yielded a product of the expected size

(1000 bp) that was then cloned using TOPO-XL vector (Invitrogen). The cloning

produced hundreds of colonies of which 96 were picked and 50 were sequenced. The

phylogenetic analysis revealed that the clone library is composed for a majority of

Micromonosporaceae (60%) and also contained gamma-, delta-proteobacteria and

Acidobacteria (Figure 4.6). One clone, OPM11 is closely related to *Micromonospora*

spp. No sequences matching exactly that of *Micromonospora* sp. strain M42 was

recovered. Using Chimera_Check, we found that seven sequences were chimeras.
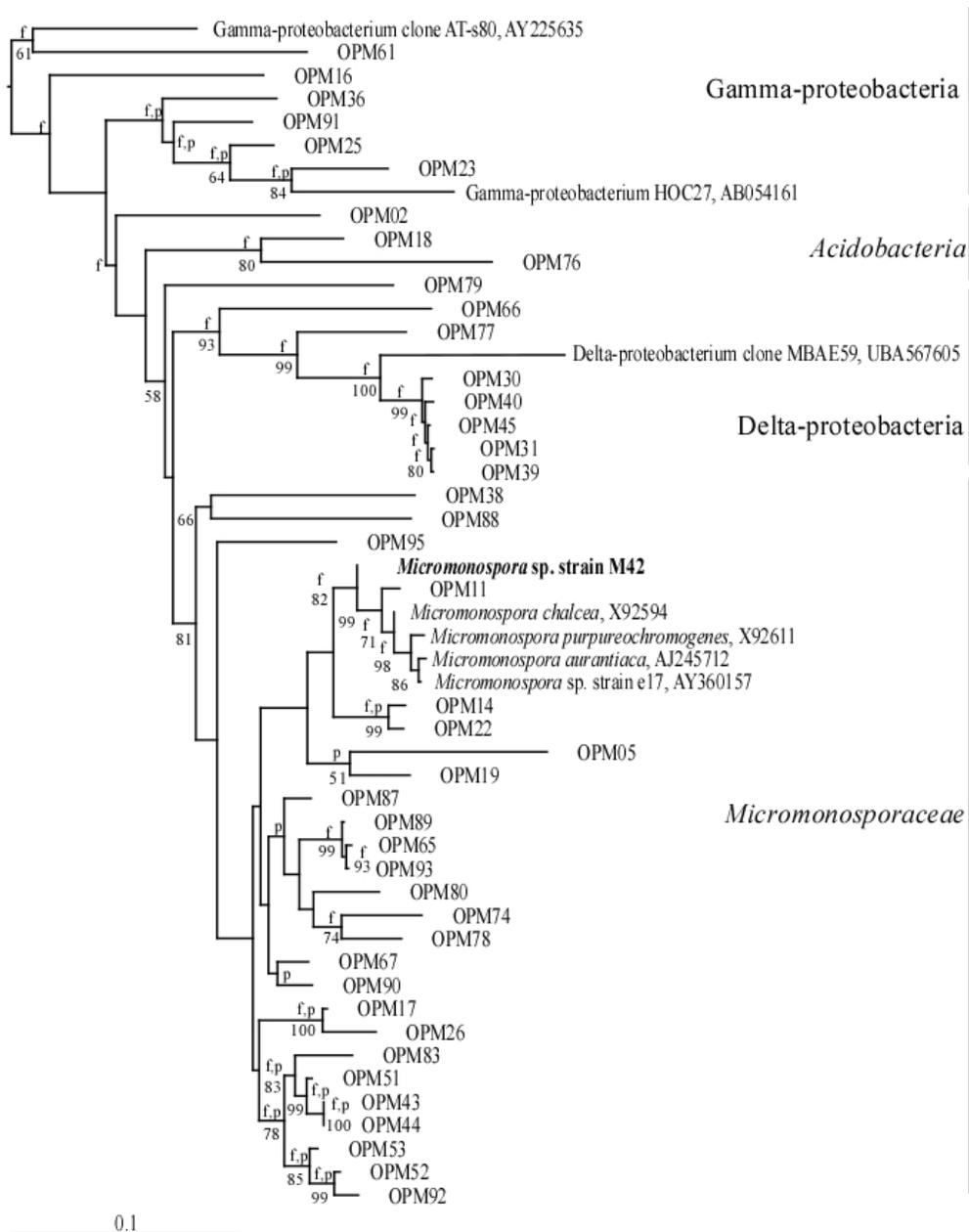
Figure 4.6. Neighbor-joining phylogenetic tree from analysis of 500 bp of 16S rRNA gene sequence of clones from the amplification of *Acanthostrongylophora* sp. Sponge 35 16S rRNA gene using primers M42-16S and 1492r. Scale bar represents 0.1 substitutions per nucleotide position.

4.3.2 Genome sizing of *Micromonospora* sp. strain M42

A preliminary investigation by PFGE showed that *Micromonospora* sp. strain M42 does not contain plasmids. The genome size of *Micromonospora* sp. strain M42 was estimated by PFGE after digesting *Micromonospora* sp. strain M42 plugs with two different restriction enzymes, *Vsp*I and *Ssp*I (Figure 4.7). Restriction digest of *Micromonospora* sp. strain M42 with *Vsp*I generated 14 fragments and digestion with *Ssp* I gave 17 fragments. The analysis of the generated restriction fragments gave an estimated genome size of 6.7 Mb.
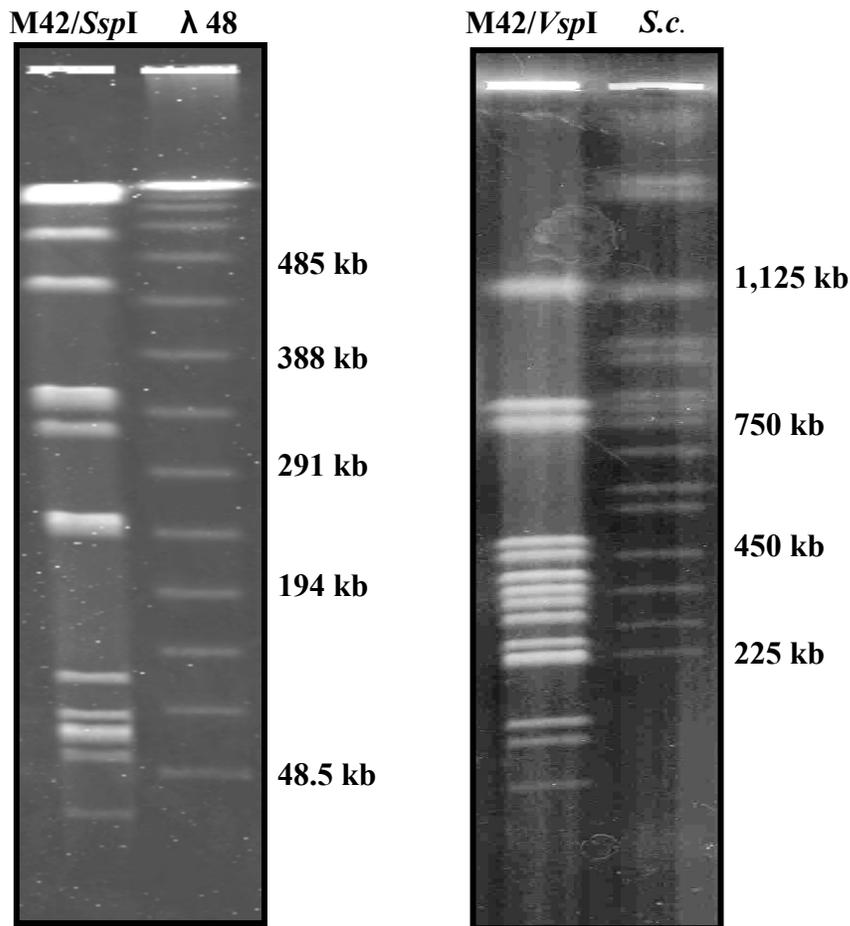
Figure 4.7. PFGE of *Micromonospora* sp. strain M42 plugs digested with the restriction enzymes *Vsp*I and *Ss* I. λ48 = Lambda ladder, *S.c.* = *Saccharomyces cerevisiae* marker.

4.3.3 Detection of polyketide synthase genes

Micromonospora sp. strain M42 was investigated for the presence of type I PKS genes. A PCR-based approach using primers MDPQQRf and HGTGTr (Kim *et al.*, 2005) yielded a 500 bp product that was cloned into a TOPO XL vector (Invitrogen). The cloning generated hundreds of colonies of which 96 were picked. The analysis of the translated sequences from eight clones revealed that seven of the sequences were identical (Figure 4.8). The only difference was observed for clone A3 for which a lysine substituted an arginine.

```
A1    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A2    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A3    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVKSRAALEHRAVVWGSDVDE
A4    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A5    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A6    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A7    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
A8    VPVPWVVSGRSERALVGQASRLASFVPGSGVSSWDVSWSLVRSRAALEHRAVVWGSDVDE
      ***********************************************:*****************

A1    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A2    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A3    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A4    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A5    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A6    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A7    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
A8    LVAGLSAVAEGRSVVSGVVSAGRRAVLFTGQGSQRVGMGRELYETFPVFAQAFDGVCAGF
      *************************************************************

A1    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A2    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A3    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A4    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A5    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A6    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A7    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
A8    EGLLPRALRDVVFAQAGSDEAALVDQTVFAQAGLFAA
      *************************************
```

Figure 4.8. Alignment of eight translated PKS sequences from clones resulting from the amplification of *Micromonospora* sp. strain M42 with primers MDPQQRf and HGTGTr. Highlighted is the only amino acid difference in the alignment, a lysine.

When compared to the GenBank database using BLASTX, all sequences were closely
related to *Streptomyces* type I polyketide synthase. The most abundant PKS sequence
was radiolabelled with ^{32}P and used as a probe for a Southern hybridization experiment.
PFGE plugs containing *Micromonospora* sp. strain M42 DNA were digested with
restriction endonucleases *Ssp*I, *Vsp*I, *Pac*I, *Dra*I and *Smi*I and the resulting fragments
were resolved by PFGE gel (Figure 4.9). The radiolabeled PKS probe hybridized to
fragments resulting from the restriction digest of *Micromonospora* sp. strain M42 plugs
confirming the presence of PKS genes throughout the *Micromonospora* sp. strain M42
genome. The plugs digested with *Ssp*I and *Vsp*I contained at least one fragment with
homology to the PKS probe.
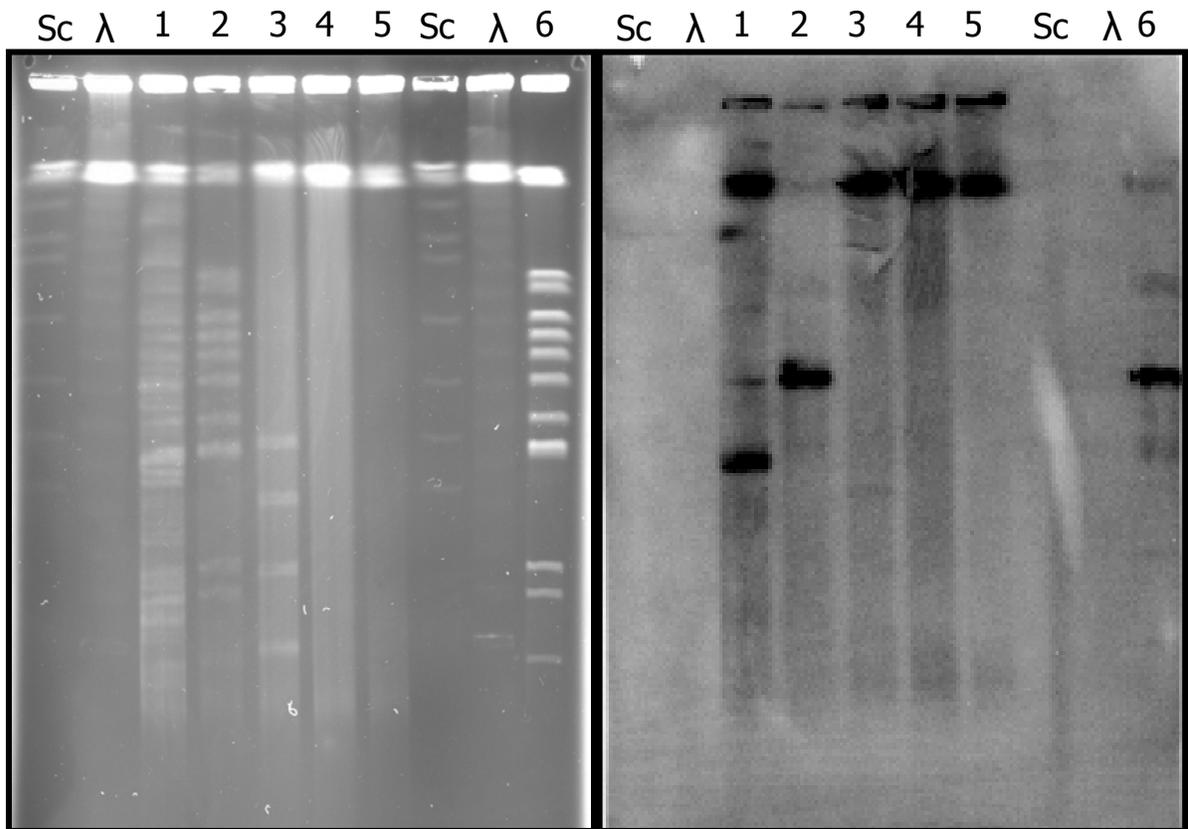
Figure 4.9. PFGE gel of *Micromonospora* sp. strain M42 plugs digested with: 1 *Ssp*I, 2
*Vsp*I, 3 *Pac*I, 4 *Dra*I, 5 *Smi*I, 6 *Vsp*I (Left). Hybridization with radiolabeled PKS probe
(Right).

4.3.4 Generation of manzamine hyper-producers by UV mutagenesis

UV mutagenesis of *Micromonospora* sp. strain M42 spores yielded 700 mutants. Of these, 300 were individually tested for improved manzamine production by thin layer chromatography (TLC) (Figure 4.10). Three mutants, M42-1, M42-1 and M42-3, showed significant improvement in manzamine production on the basis of darker spots on TLC analysis. Further testing for quantification purposes using LC-MS-TOF was carried out and revealed that M42-1 produced 3.5 mg/l of manzamine A, M42-2 produced 2.95 mg/l of manzamine A and M42-3 produced 1.63 mg/l of manzamine A. Additionally, one of the mutant, M42-2, produces manzamine B in a yield of 4 mg/l which is the postulated precursor of manzamine A (Hu *et al.*, 2003).
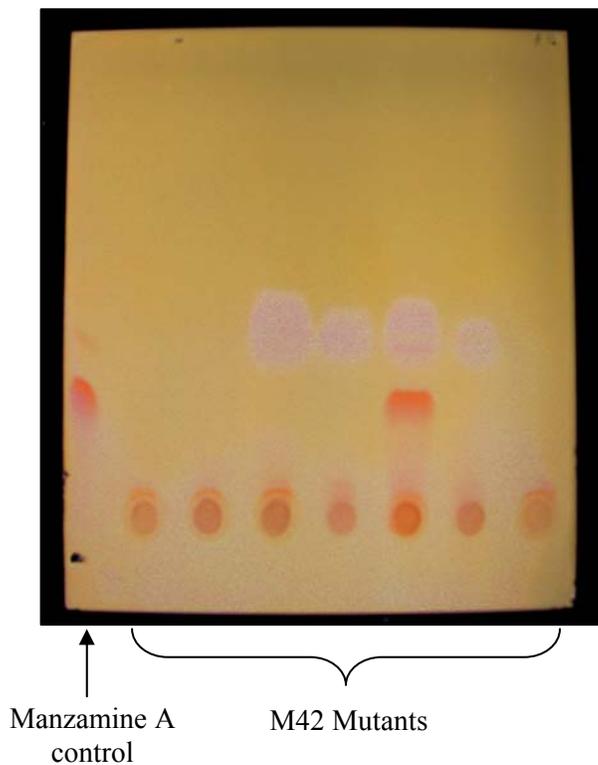
Figure 4.10. Thin layer chromatography analysis of *Micromonospora* sp. strain M42 mutants.

4.3.5 Manzamine A pathway

Two approaches were used in an attempt to identify the genes involved in the biosynthesis of manzamines. Both approaches were based on the hypothesis that a tryptophan decarboxylase enzyme is produced by *Micromonospora* sp. strain M42 to catalyze the decarboxylation of tryptophan to obtain tryptamine, which is part of the beta-carboline moiety of the manzamine molecule. The molecular approach using six different sets of PCR primers (Table 4.3) in different arrangements yielded three products. The primers pair TDC1f/TDC3r gave a 450 bp product, TDC1f/TDC4r yielded a 180 bp product and TDC2f/TDC3r gave a 210 bp product. The sequence analysis and the comparison to GenBank database using BLAST revealed that all the amplified products were closely related to the *Streptomyces griseus* secA gene. Products of secA genes are membrane proteins involved in the translocation of preprotein.

	Primers	Sequence (5'-3')
Forward	TDC10	TGGATGCACACGUCGACGG
	TDC11	TGGATGCACACGTGGACGG
	TDC12	CATCCACAGCTTCAGGGC
Reverse	TDC13	CATCCACAGCTTCAGCGC
	TDC14	CGGAAGACCAGGGT
	TDC15	CGGAACACCAGGGT

Table 4.1. Sequence of tryptophan decarboxylase primers

In a biochemistry approach to the detection and isolation of tryptophan decarboxylase from *Micromonospora* sp. strain M42. Total protein was extracted from cell material. The activity of *Micromonospora* sp. strain M42 protein extracts was assessed by measuring the transformation rate of tryptophan into tryptamine using HPLC. Extracts at

various stages of purification, including original extract, first ammonium sulfate treatment extract and DEAE-Sephadex fractionated extracts were tested. No tryptophan decarboxylase activity could be detected.

4.4 Discussion

Assessing the growth of actinomycetes in liquid culture can be challenging because of the mycelia growth form of these bacteria. Formation of extensive mycelia in some actinomycetes can make the use of optical density problematic leaving the tedious method of plate counts and dry weight determinations as the only choice to assess the growth. The results of the comparison between OD_{600} and plate counts to monitor the growth of *Micromonospora* sp. strain M42 showed that both methods generated very similar growth curves. Though *Micromonospora* sp. strain M42 cells have mycelia, microscopic examination revealed that this was limited (Figure 4.11).
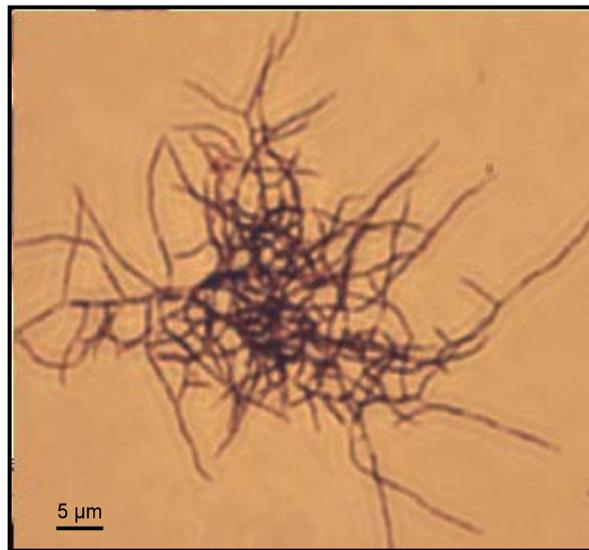
Figure 4.11. Micrograph of Gram stained *Micromonospora* sp. strain M42 mycelia.

The extent of mycelial formation did not seem to interfere with the use of optical density measure of growth. The advantages of OD_{600} over plate counts are the practicality of the method as well as the rapidity for obtaining the results which are obtained instantly whereas the plate counts methods requires an incubation time of seven days before obtaining results. Another advantage of OD_{600} is that many fermenters are fitted to monitor optical density, making it more practical in scaled up fermentations to monitor the growth of the organism and evaluate the best time to harvest the culture.

During a 16 day time course experiment where both manzamine A production and the growth of *Micromonospora* sp. strain M42 were monitored. Variations during the fermentation experiments were observed and are likely due to the size of the inocculum, the aeration of the media and the shaking of the flasks. Manzamine A could be detected after just one day after inoculation and reached a maximum in the middle of the exponential phase at day 7. If manzamine A was a secondary metabolite, a compound which is not required for the normal growth or development of the organism, it would be expected that manzamine A would be produced during the stationary phase. The fact that manzamine A can be detected very early during the growth until the mid log phase points toward a constitutive expression of the metabolite. It is possible that because of the very specific ecological niche in which *Micromonospora* sp. strain M42 lives, the organism has been under ecological pressures that triggered a shift for manzamine A from a secondary metabolite type expression to a constitutive and necessary expression. Although the bioactivities of manzamine A have been well documented (Hu *et al.*, 2003), the ecological role of manzamine A is unknown. It can be speculated, in line with the

known toxicity of manzamine A and the fact that it is excreted into the medium, that *Micromonospora* sp. strain M42 produces the compound as a defense mechanism.

Micromonospora sp. strain M42, despite its isolation from a marine invertebrate, achieves better growth with little (1% NaCl) or no salt in the medium. These results are consistent with previous observations for the growth of *Micromonospora* isolated from marine sediments where optimum growth was observed at about 1% NaCl (Weyland, 1981). The amount of manzamine produced by *Micromonospora* sp. strain M42 was found to be higher with no or little NaCl (1%) in the medium matching the cell density results. Interestingly, the ratio OD_{600} to manzamine A shows that more manzamine A is produced by less cells at higher NaCl concentrations, 3.5% and 5% NaCl, this could possibly explain why with a limited amount of manzamine A-producing bacteria in the sponge, the manzamine A concentration within the sponge tissue can reach high levels. It is unclear what the salt concentration level is in the sponge mesohyl is. One study carried on *Microciona prolifera* (Philp, 2001) reported that sea salt concentration and pH have an influence on sponge cell Cd^{2+} uptake and their faculty to aggregate.

Media used for the isolation of sponge-associated bacteria are routinely supplemented with 2% NaCl assuming that bacteria associated with marine invertebrates have requirements for a higher salt concentration than terrestrial microbes. The isolation of microbes is often a race won by the faster growing organisms. The results from the salinity experiments showed that *Micromonospora* sp. strain M42 grew significantly faster with a lesser salt concentration than the 2% NaCl used. With the objective of

131

isolating as many bacteria as possible, the use of media with lower salt concentrations may yield isolates of additional sponge-associated bacteria that would be overgrown by other marine bacteria on media with higher NaCl concentrations.

For large scale production of *Micromonospora* sp. strain M42 in bio-reactors, the limited NaCl requirement of the strain is an advantage. Use of low-salt media will prevent fermenters from potential corrosion that may occur with higher NaCl concentrations.

The cloning of the PCR product resulting from the amplification of *Acanthostrongylophora* microbial 16S rRNA gene with primers M42-16S and 1492r showed a greater diversity of bacteria than expected, as originally primer M42-16S was designed to amplify *Micromonospora* sp. strain M42. The clone library contained a minority of unexpected groups including gamma- and delta-proteobacteria, *Acidobacteria* and a majority of *Micromonosporaceae*. It is likely that the M42-16S primer has a sequence that matches or is a close match to that of the amplified organisms. Several chimeric sequences were found in the clone library. The formation of chimera occurs during the PCR process, when DNA molecules with high sequence similarity compete with specific primers during the annealing step. The formation of chimeras can be reduced by increasing the annealing time (Von Wintzingerode *et al.*, 1997). Also the pairing of M42-16S with eubacterial primer 1492r may have limited the specificity of the PCR amplification. It is likely that the pairing of M42-16S with a different primer more specific to actinomycetes or the *Micromonospora* genus, would improve the specificity of amplification of *Micromonospora* spp. Nevertheless, the large majority of organisms in the library belonged to the *Micromonosporaceae* and one was very closely related to

Micromonospora sp. strain M42. Interestingly, the nested PCR amplification, in which

one amplification reaction was with a *Micromonospora* "specific" primer, showed a

diversity of *Micromonosporaceae* that was not previously revealed by the initial

Acanthostrongylophora 16S rRNA gene clone library (Figure 2.1). The diversity

Micromonosporaceae offers a possible explanation for the large amount of manzamine A

found in the sponge. Though *Micromonospora* sp. strain M42 is the only manzamine A

producer that has been isolated, it is possible that the manzamine A production

capabilities are shared with closely-related bacteria. Horizontal gene transfer events may

have occurred between closely related *Micromonospora* strains clustered together within

the sponge tissue, resulting in the spread of the manzamine biosynthetic capability within

this assemblage.

The genome size of *Micromonospora* sp. strain M42 is estimated at 6.7 Mb. It is a rather

large genome compared to non-actinomycete bacteria though much smaller than other

actinomycetes such as *S. coelicolor* (9.05 Mb) or *S. avermitilis* (9.12 Mb). These two

actinomycetes are extremely prolific secondary metabolite producers. Symbiotic

bacteria live in very specific and controlled environments and in some cases, symbionts

have been found to have smaller genomes than their non-symbiotic counterparts. Based

on the assumption that the genome size of a symbiont can be an indicator of the duration

of the association between the symbiont and its host, the genome size of

Micromonospora sp. strain M42 suggests that its association with its sponge host could

be relatively recent.

Actinomycetes are known to produce multiple secondary metabolites and genomic

studies permitted the prediction of metabolites that were not detected before (Challis &

Ravel, 2000). *Micromonospora* sp. strain M42 was investigated for the presence of type

I polyketide synthases (PKSs) and was found to have at least two different PKSs. It is

likely that *Micromonospora* sp. strain M42 contains more secondary metabolite genes

including PKS and non-ribosomal peptide synthases (NRPSs) in its genome. The

analysis of *Micromonospora* sp. strain M42 cultures was focused exclusively on

detecting alkaloids and more specifically manzamines, therefore it is likely that other

secondary metabolites would not be detected. The evidence of PKS genes strongly

encourages future searches for this type of metabolites.

Mutation is one of the key methods for the improvement of microbial strains producing

bioactive compounds. Mutagenesis can be induced chemically using colchicine or

ethylmethane sulfate or by exposure to UV light and can result in a 10- to 1000- fold

increase in the production of metabolites (Adrio & Demain, 2006). Mutagenesis of

Micromonospora sp. strain M42 yielded three hyper-producing mutants that had

improved manzamine A production ranging from a 2.3 to 3.5 fold increase. Interestingly,

one of the mutant strains, M42-2 was shown to produce manzamine B (Figure 4.12).
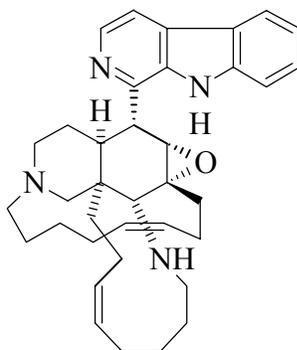
Figure 4.12. Manzamine B

Manzamine B is the proposed precursor to manzamine A but it is also a better molecule

than manzamine A for generation of analogs. The three mutants obtained are an

important step toward obtaining strains producing high yields of manzamines. They will

undergo further mutations to improve the manzamine production levels.

A biogenetic path for manzamine A has been proposed (Baldwin & Whitehead, 1992)

(Figure 4.12) it involves an intramolecular Diels-Alder reaction of a tricyclic ring system

(*bis*-3-alkyldihydropyridine) containing a dihydropyridine (the diene) and a conjugated

iminium ion (the dienophile), followed by disproportionation and hydrolysis of the

intermediate iminium ion to give a tetracyclic aldehyde (reminiscent of ircinal B) which

was subsequently reacted with tryptophan to yield manzamine B, followed by manzamine

A.
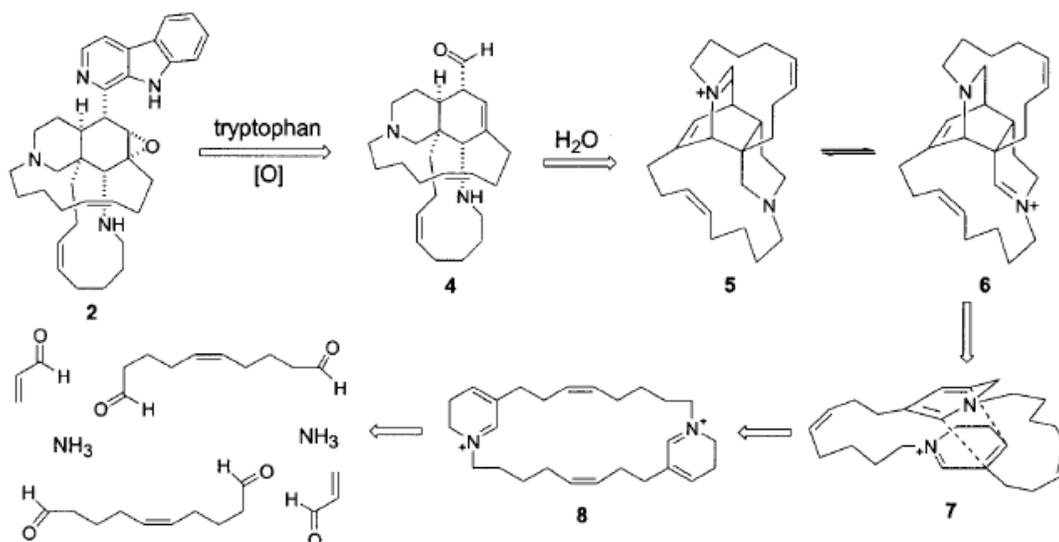
Figure 4.13. Baldwin and Whitehead's hypothesis for the biosynthesis of the manzamine alkaloids.

The chemistry of manzamines is relatively well known but the biogenesis of these compounds remains unclear. We focused on the tryptophan decarboxylase enzyme, making the assumption that a tryptophan unit undergoing decarboxylation was involved in the formation of the beta-carboline group. Neither molecular or biochemistry approach was successful in detecting the presence of a tryptophan decarboxylase in *Micromonospora* sp. strain M42. As the biosynthesis of manzamine A is speculative, it is possible that no tryptophan unit is involved in the synthesis of manzamine and a tryptophan decarboxylase enzyme is absent. In the event that a tryptophan decarboxylase is involved in manzamine A synthesis, it is possible that the molecular approach failed because of a lack of homology between the enzymes used in design of the PCR primers and the actual sequence of *Micromonospora* sp. strain M42 tryptophan decarboxylase gene. For the biochemistry approach, the tryptophan decarboxylase enzyme could be unstable or require specific cofactors that made it difficult to detect the activity of the enzyme.

The isolation of the manzamine producer *Micromonospora* sp. strain M42 proves that manzamines can be of bacterial origin as hypothesized by Kobayashi and Ishibashi (1993). It provides a working platform for the study of manzamines which have been regarded as an intriguing group of marine alkaloids with extraordinary biological activity. There is no doubt that the mystery of the biosynthetic pathway of manzamine A will be elucidated in the near future and that *Micromonospora* sp. strain M42 will be the major contributor to that discovery.

5 Low coverage sequencing and analysis of the *Micromonospora* sp. strain M42 genome

5.1 Introduction

5.1.1 Whole genome shotgun sequencing.

In 1995, The Institute for Genomic Research led by J. Craig Venter, completed the first entire genome of a free living organism, *Haemophilus influenzae,* by using whole genome shotgun sequencing (Fleischmann *et al.*, 1995). Since that first report, more than 400 microbial genomes have been completely sequenced and at least another 932 are in progress (July 2006 - http://www.genomesonline.org/). This global effort has focused primarily on pathogens, which to date account for the majority of all genome projects, and has generated a large amount of raw material for *in silico* analysis. In recent years, multiple strains of the same species or multiple species of the same genus have been the targets of sequencing projects, opening the possibility of comparing closely related genomes. This will improve our understanding of microbial biology, pathogenicity and evolution. The major challenge in the post-genomic era remains to fully exploit and decipher this new accumulating wealth of information.

The whole genome shotgun sequencing approach involves the breaking of large DNA molecules into smaller pieces that are then cloned into vectors such as plasmids for insert size up to 12 kb) or fosmids for larger insert up to 45 kb. The sequencing of these

randomly cloned inserts provides overlapping sequence reads. The whole genome

shotgun sequencing strategy does not require an initial mapping step to create a set of

overlapping clones, and instead relies on computational methods, TIGR Assembler

(Sutton *et al.*, 1995), the Celera Assembler (Myers *et al.*, 2000) or Phrap

(http://www.phrap.org), to correctly assemble tens of thousands of random DNA

sequences 300-900 bp long.

Whole genome shotgun sequencing can be broken down in six steps (Figure 5.1). First a

library is constructed from total genomic DNA extracted and sheared into smaller

fragments that are then ligated into cloning vectors. The next step consists in randomly

sequencing the clones. Enough clones to achieve 8X coverage are sequenced which

amounts to ca. 6000 random clones /Mb. The third step assembles all the small

sequences (~ 800 bp) into larger contigs using computational algorithms such as the

Celera Assembler. In the following step, the contigs are linked to each other during the

closure phase, where the sequence is also manually edited. Then programs such as

Glimmer identify open reading frames (step 5) and the predicted protein sequences from

these putative open reading frames are searched against non-redundant protein databases.

Finally, the last step is the obtention of the complete genome which is obtained after

manual curation of the annotation.

The analysis of a completed and fully assembled genome begins by determining the

precise location and assigning a putative function to all the protein coding regions,

through a process known as annotation. A wide variety of bioinformatics methods that

have been developed to analyze sequence data have made annotation an increasingly

sophisticated process. Computational gene finders using Interpolated Markov modeling

algorithms such as Glimmer (Delcher *et al.*, 1999) are routinely capable of finding more than 99% of all genes in a microbial genome. The predicted protein sequences from these putative open reading frames (ORFs) are searched against non-redundant protein databases and well-curated protein families such as the PFAM (Bateman *et al.*, 2002) and TIGRFAM (Haft *et al.*, 2003) collections that have been created using Hidden Markov Models (HMMs). HMMs are powerful statistical representations of groups of proteins which share sequence, and consequently, functional similarity. HMMs can represent very specific enzymatic functions or a superfamily of related functions. The use of HMMs has helped refine the annotation process. In addition, searches for PROSITE motifs (Sigrist *et al.*, 2002), lipoproteins, signal peptides and membrane spanning regions are performed. Based on the evidence gathered, a two-stage annotation protocol is carried out whereby an initial automated annotation is followed by manual curation of each gene assignment by an expert biologist to ensure accuracy and consistency of the putative function of each predicted coding region. Proteins whose specific function cannot be confidently determined are designated "putative" or given a less specific family name. Proteins without any significant matches in any of the searches are annotated as hypothetical proteins. Consistent description and annotation of genes in different databases is critical to facilitate uniform queries across independent databases.

The Gene Ontology (GO) project ensures that a normalized vocabulary is used for the description of gene products (Gene Ontology, 2004). Following GO guidelines, gene products are in terms of their associated biological process, cellular components and molecular functions in a species independent manner.
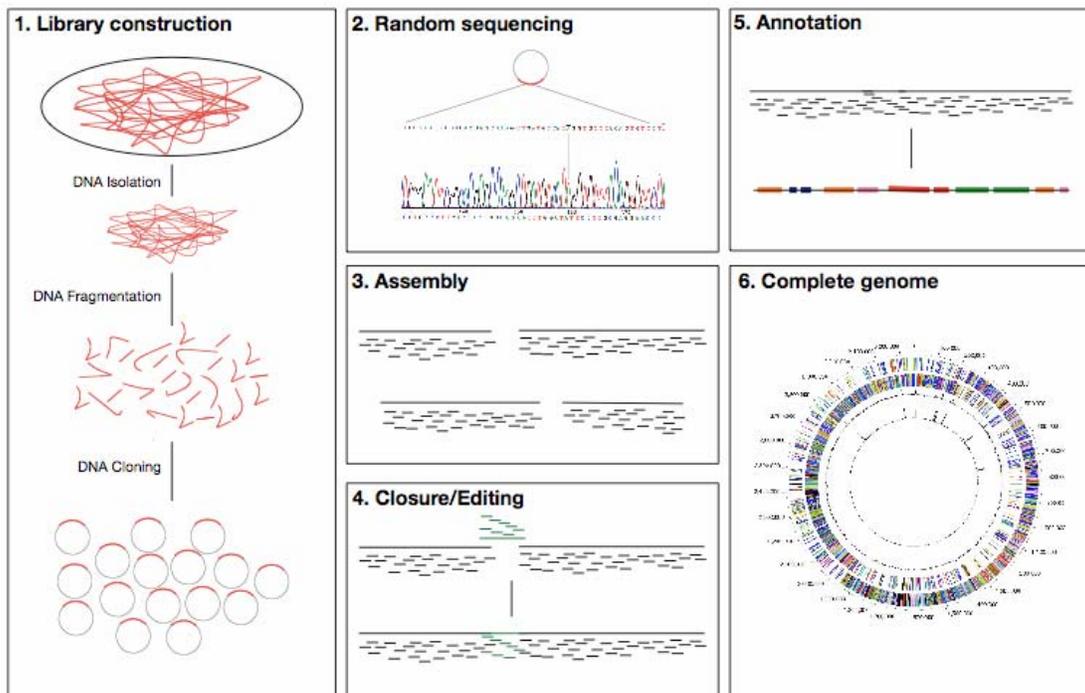
Figure 5.1. Steps involved in the whole genome shotgun sequencing procedure.

5.1.2 Contributions from genomics

High-throughput genome sequencing technologies have only been around for about 10 years, but the impact of these technologies has been profound. Genome sequence data have been obtained from representative species of all three domains of life, however, because of their relatively small size, bacterial and archaeal genomes have dominated the field. Taken together, comparative genome analysis has revealed interesting patterns pertaining to microbial species. For example, gene density in microbes is very consistent with about one gene per kilobase of DNA. Although we are

able to identify microbial genes with a high degree of success, we cannot assign a function to about a quarter of all the ORFs in each species sequenced so far. This observation demonstrates how little is known about the biology and biochemistry of microbial species, and supports the idea of an incredible microbial diversity. These sets of genes that encode hypothetical proteins represent exciting opportunities for the research community and are potential sources of biological resources to be explored for future use, but also clearly indicate the need for further extensive genetic, enzymatic and physiological analyses before genomic data can be fully exploited.

Analysis of more than 150 microbial genome sequences has revealed an unexpected diversity and variability in genome size and structure, even in species previously thought to be identical. Many microbes possess diverse chromosome architectures quite different from the classical single circular chromosome. For example, the genome sequence of the human pathogen, *Vibrio cholerae*, unexpectedly revealed the presence of two circular chromosomes (Heidelberg *et al.*, 2000), whereas the genome of *Borrelia burgdorferi* (Casjens *et al.*, 2000; Fraser *et al.*, 1997), the causative agent of Lyme disease, contained a relatively small (910 kb) linear chromosome and an unprecedented number of 21 linear and circular plasmids. On the other hand, the *S. coelicolor* linear chromosome is more than 9 Mbp long (Bentley *et al.*, 2002). In addition to differences in genome structure, microbial genomes vary largely in their GC content ranging from 24% to more than 70%. The effect of this disparity in GC content is reflected in the wide range of codon usage and the amino acid composition of proteins among various species.

Organisms that belong to the same genus can differ in gene content by as much as 25% as it was found when the genome of *E. coli* K-12 was compared to *E. coli* 0157:H7 (Hayashi

et al., 2001). Insertion and deletion events appear to have played a major role and account for most of the differences observed. Pathogenicity islands, which are large blocks of self-mobile DNA that carry genes enabling an organism to act as a pathogen, have the ability to transfer from one organism and integrate into a new host. Other pathogens show little variation in chromosomal gene content, as demonstrated by the comparison of the genomes of two isolates of *Yersinia pestis* (Deng *et al.*, 2002; Parkhill *et al.*, 2000), the etiologic agent of plague. Remarkable differences in the chromosome structures, dominated by genome rearrangements, accounted for most of the variation observed between these two closely related strains. The differences appear to result from multiple inversions of genome segments at insertion sequences. *Y. pestis* sp. carry most of their virulence determinant on plasmids, which are absent in its ancestor, *Yersinia pseudotuberculosis*. A remarkable number of pseudogenes (degenerated and inactive genes) have been found on the genomes of *Y. pestis*, an indication of a recent and still evolving genome.

The initial focus on pathogenic microbial species has shifted to include non-pathogenic environmental microbes. Understanding and accessing the tremendous microbial biochemical diversity that exists in the environment could have an important impact on industrial processes and help in resolving environmental issues such as the bioremediation of human pollution.

Many archaea are considered extremophiles, as they often thrive under "extreme" conditions, such as high or low temperatures, high pressures or high salt concentrations among others. The novel enzymes encoded in these genomes offer clear potential for biotechnological applications. In addition, genome analysis of the hyperthermophilic

bacteria, *Thermotoga maritima* (Nelson *et al.*, 1999) revealed that 20-25% of the genes in this species were more similar to genes from archaea than from bacteria, leading to a renewed interest in the process of lateral gene transfer and the role that it plays in microbial evolution and diversity.

Among the bacteria, the genome sequence of *Photorhabdus luminescens* (Duchaud *et al.*, 2003), an insect-pathogen living in symbiosis with a nematode revealed a variety of genes coding for entomopathogenic toxins, potentially useful in defending against insect pests. *P.* luminescens carries also a large number of genes coding for the biosynthesis of antibiotics and fungicides, which could have potential applications for the treatment of infectious diseases. In 2002 and 2003, the genomes of two industrially important filamentous actinomycetes, *S. coelicolor* (Bentley *et al.*, 2002) and *S. avermitilis* (Ikeda *et al.*, 2003; Omura *et al.*, 2001) were completed. These strains are known to participate in the degradation and recycling of complex organic polymers in soil. Most important is their ability to produce a diverse array of secondary metabolites with bioactive properties including the calcium dependent antibiotic (CDA) (Hopwood & Wright, 1983) from *S. coelicolor* and the anthelmintic avermictin (Burg *et al.*, 1979) from *S. avermitilis*. The genome sequence of these two actinomycetes revealed the strikingly large number of gene clusters devoted to the synthesis of secondary metabolites, 23 gene clusters for *S. coelicolor* and 30 for *S. avermitilis*. Access to the genome sequence of these representative strains will assist in genome engineering to make novel and more efficient antimicrobial agents.

Researchers have only scratched the surface of microbial biodiversity. In order to harvest this enormous potential, genome shotgun sequencing is being applied to the environment.

In a landmark study, the microbial populations from water samples collected in the Sargasso Sea were sequenced (Venter *et al.*, 2004). An estimated 1.2 million new genes have been identified from at least 1,800 genomic species. Similar techniques were applied to a community of microbes from a biofilm growing at pH 0.83 on the surface of acid mine drainage (Tyson *et al.*, 2004). In this study, the low diversity genomic community was entirely reconstructed and the subsequent examination of the metabolic capabilities of this community gave valuable information on how each organism participates to the ecology of the biofilm. These types of microbial studies will help to define the entire repertoire of organisms in specialized niches and the mechanisms by which they interact with their environment.

Microbial genome sequencing provides an unrestricted access to the entire genetic potential of an organism and contributes to a better understanding of the microbe. However, the sequence leaves major gaps with many identified protein for which no functions can be assigned (Boucher *et al.*, 2001).

For pathogenic agents, genome sequence analysis has provided critical information such as potential new drug targets and sequence target for rapid PCR-based diagnostic tools (Rosamond & Allsop, 2000). For microbes that produce bioactive compounds such as *Streptomyces* and *Micromonospora* sp. strain M42, genome sequencing assists in establishing a detailed map of their primary and secondary metabolism. Primary metabolisms pathways are important because they are often the suppliers of precursor used as building blocks for the bioactive compounds. Knowledge of these pathways can help in designing media that will ensure a steady supply of precursor therefore optimizing the production of the bioactive compound of interest (Donadio *et al.*, 2002).

The importance of sequencing *Micromonospora* sp. strain M42 is two-fold. Because of *Micromonospora* sp. strain M42's ability to produce manzamines, analysis of its genome sequence is expected to reveal the manzamine's biosynthetic gene cluster and to establish *Micromonospora* sp. strain M42's secondary metabolism potential. The second reason for sequencing the genome of *Micromonospora* sp. strain M42 lies in the role of *Micromonospora* sp. strain M42 as a sponge symbiont. No sponge-associated bacteria have been sequenced and little is known about the effect of the sponge ecological niche on bacteria. The mechanism whereby sponges select and acquire their bacterial symbionts remains unclear. It has been proposed that bacteria resisting the sponge digestion process would have the potential to become symbionts (Wilkinson *et al.*, 1984) but hypotheses involving molecular interactions were not proposed. The plant symbiont *Mesorhizobium loti* possesses genetic symbiosis islands that permit its symbiotic relationship with its lotus host and provide the host with nitrogen (Sullivan *et al.*, 2002). It remains to be determined if for sponge-associated bacteria, some molecular factors such as symbiosis islands are involved in the establishment of sponge/bacteria relationships.

The low-coverage (1X) genome sequencing of the *Micromonospora* sp. strain M42 genome is a first step to a better understanding of the bacterium's biology. By focusing on some of the key functions of the microbe, including transporter system, amino-acid synthesis, two-component regulatory system, secondary metabolite capabilities and phages and mobile elements, valuable information can be gathered that will improve our understanding the ecology and biology of this strain and will provide insights into the complex relationship between *Micromonospora* sp. strain M42 and the

Acanthostrongylophora sponge. In addition, the genomic information will open the

possibility to rationally engineer this strain to improve manzamine production.

5.2 Materials and Methods

5.2.1 Culture and DNA extraction

Micromonospora sp. strain M42 was inoculated in 50 ml of ISP2 medium in a

250 ml flask. The culture was incubated at 30˚C and shaken at 180 rpm for 5 days. DNA

was extracted following the modified method of Hopwood *et al.* (1985) (Appendix 4).

To limit DNA shearing, large bore tips were used and DNA was spooled after being

precipitated. The size of the DNA extracted from *Micromonospora* sp. strain M42 was

assessed by gel electrophoresis on a 0.8% agarose and shown to be around 40 kb (Figure

5.2). *Micromonospora* sp. strain M42 DNA sample was purified using Genomic-Tip kit

(Qiagen) according to the manufacturer's recommendations.
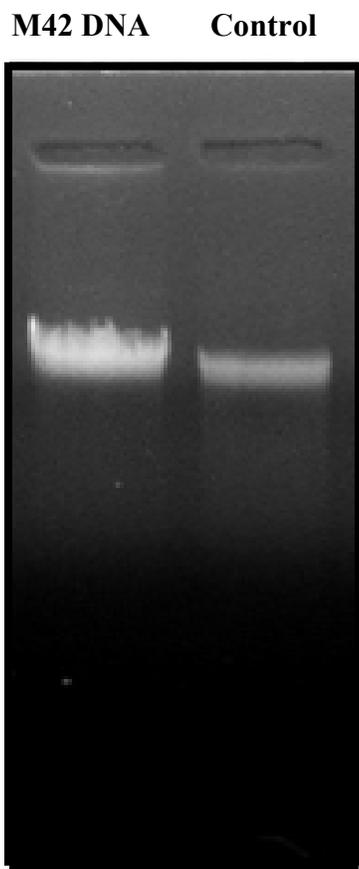
Figure 5.2. Electrophoresis gel image of *Micromonospora* sp. strain M42 extracted DNA (M42 DNA) and Epicentre control DNA (36 kb). *Micromonospora* sp. strain M42 sheared DNA size is estimated to be around 40 kb.

5.2.2 Cloning and sequencing of *Micromonospora* sp. strain M42

Purified *Micromonospora* sp. strain M42 genomic DNA was sheared by pipeting three times up and down in 100 µl tips, end-repaired and 30-45 kb fragment were ligated into fosmid pCC1FOS vector (Epicentre Technologies, Madison, WI) following the manufacturer's instructions. Fosmid clones (4,600) were sequenced from both end on a ABI3730xl sequencer giving a total of 7,568 good reads (82% success rate) with an average read length of 618 bp for a total of 4.7 Mb of sequence.

5.2.3 Bioinformatic analysis of *Micromonospora* sp. strain M42 fosmid
library

The 7,568 sequences were compared to GenBank and TIGR databases (OMNIUM: a curated non-redundant database comprising of all sequenced genomes) using BLASTX (Altschul *et al.*, 1990). The assigned threshold level for hit significance was set to 1e-3. The E-value or Expect value describes the random background noise that exists for matches between sequences. The lower the E-value, the more significant the match. The top BLASTX hit was collected for each sequence read and stored. The data was parsed using custom-design perl scripts. The header of each entry in OMNIUM contains if available, the following information related to the entry: organism name, common name, gene symbol, EC number and functional role category. EC number is the enzyme commission number, a numerical classification for enzymes based on the chemical reactions they catalyze. An example of the results from BLASTX for the sequences discussed in this chapter is shown in Appendix 5. The functional role categories are represented by one of the 102 numbers adapted from Riley (1993), and is used at TIGR for the classification of cellular functions.

5.3 Results and Discussion

5.3.1 Overview

A total of 7,568 reads were obtained from 35-45 kb insert size fosmid library of *Micromonospora* sp. strain M42. With an estimated genome size of 6.7 Mb (Chapter 4) and an average read size of 618 bp, these 7,568 reads represent 0.7 fold sequence coverage of the genome. Using BLASTX, the reads were compared to the GenBank and TIGR databases for homology with existing identified proteins. The general results from this analysis (Figure 5.3) showed that the majority of identified sequences (636 sequences) were homologous to genes involved in energy metabolism (Figure 5.3). Other well-represented functions include transport and binding proteins, cell envelope, regulatory functions, protein fate, central intermediary metabolism and cellular processes. Functions with a smaller representation included protein synthesis, DNA metabolism, biosynthesis of cofactors, prosthetic groups and carriers, amino acid biosynthesis, fatty acid and phospholipid metabolism, transcription, purines, pyrimidines, nucleosides and nucleotides, mobile and extrachromosomal element functions and signal transduction. Sequences representing close to one fourth of the total number of sequences (24%), were classified as hypothetical proteins or having unknown functions.
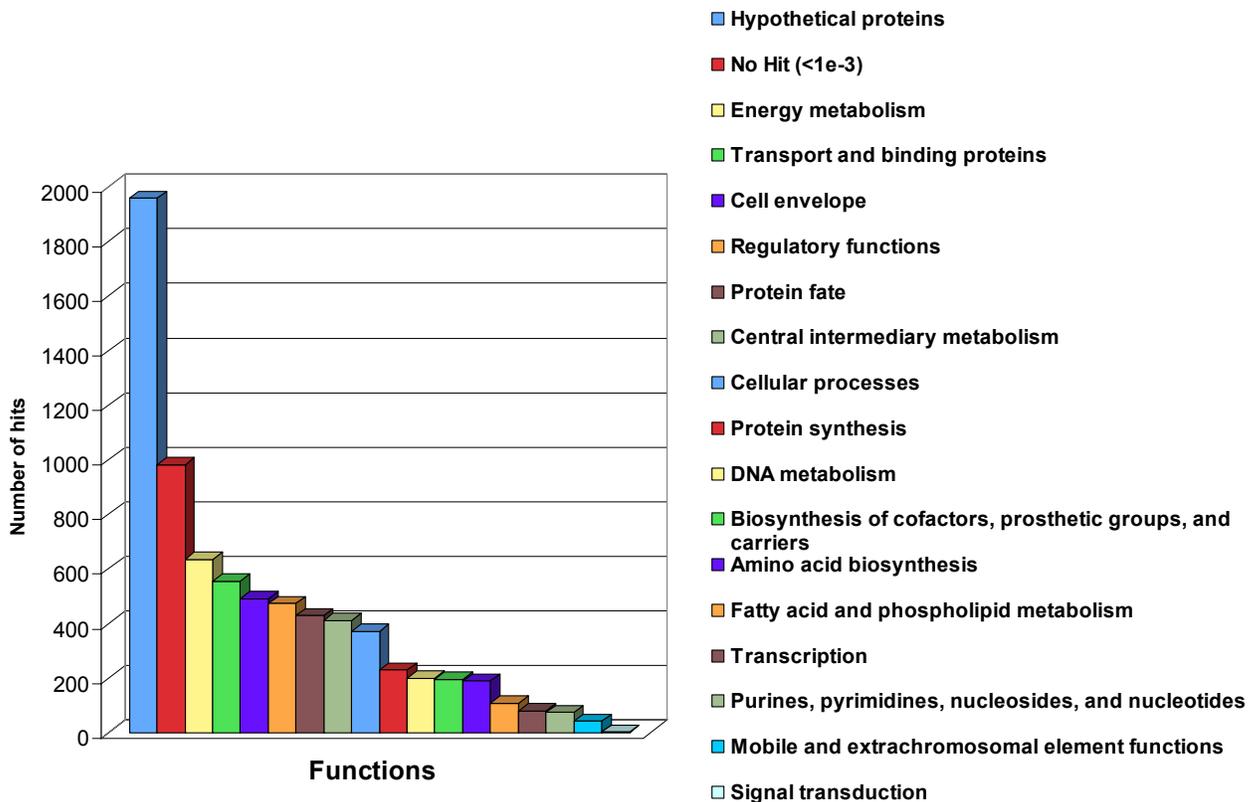
Figure 5.3. Best hit distribution by functional category

A number of sequences accounting for 13% had no meaningful homology (E value <1e-3) to proteins in the databases.

The comparison of the 7,568 sequences with the protein database showed that 77% of the top BLASTX hits were to other actinomycete proteins including proteins from *Streptomyces clavuligerus (*47%), *Mycobacterium smegmatis* strain MC2 (11%), *Arthrobacter aurescens* (8%), *Mycobacterium avium* (7%), *Mycobacterium tuberculosis* (3%) and *Actinomyces naeslundii* strain MG1 (2%) (Figure 5.4). Top BLASTX hits to proteins from 61 other organisms accounted for less than 15% of the total hits.
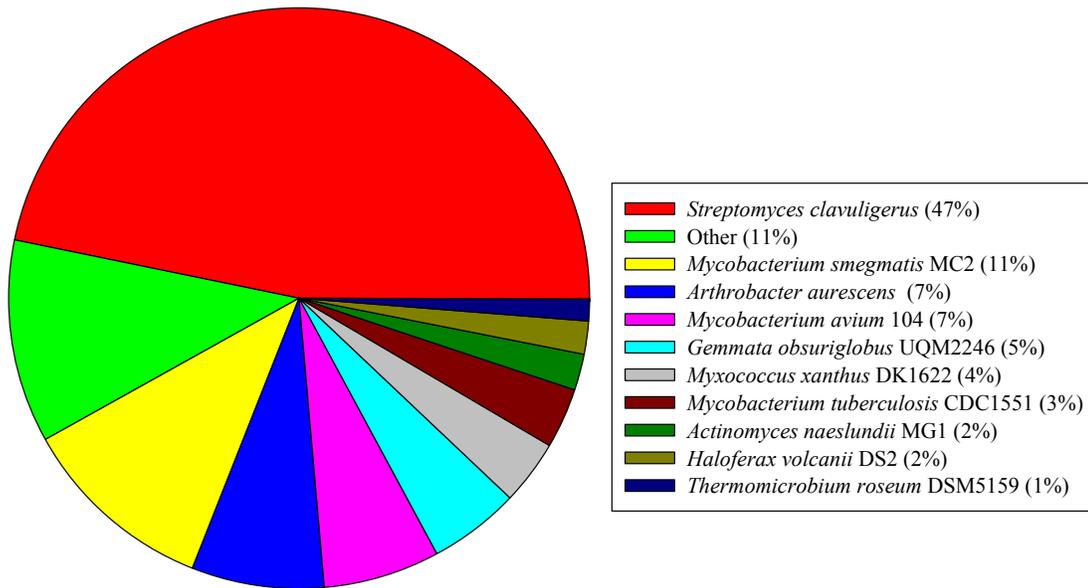
Figure 5.4. Distribution of top BLASTX hits from *Micromonospora* sp. strain M42 sequences according to the organisms from which the hits come. Number in brackets represents the percentage of total sequences reads that hit the organism. "Others" includes 61 organisms which accounted for 11% of the hits.

5.3.2 Transporter systems

The number of sequences encoding for proteins linked to transport function (524) represents 7% of the total number of sequences. The majority of these sequences (54%) are identified as being part of the ABC transporter system. The ATP-binding cassette (ABC) transporters are one of the largest families of transport proteins and accounts for about 5% of the genome in *E. coli* (Blattner *et al.*, 1997). It is a ubiquitous system that can be found in archaea, eukaryotes and eubacteria (Higgins 1992). In Gram-negatives, it is known as the periplasmic binding protein-dependent transport system and in Gram-positive bacteria as the binding-lipoprotein-dependent transport system. A typical

organization of the ABC transport system consists of two permeases each having six transmembrane segments, two peripheral membrane proteins that bind and hydrolyze ATP, and a periplasmic susbstrate binding lipoprotein. The ABC transporters form the largest group of paralogous genes in bacterial and archaeal genomes and genes are frequently organized in an operon (Higgins, 1992). Elements of the ABC transport system has been shown to mediate multi-drug resistance for *Aspergillus nidulans* (Andrade *et al.*, 2000) and a recent study reported that an ABC transporter associated with a membrane protein from the actinomycete *Planobispora rosea* was able to stimulate antibacterial activity in *Streptomyces lividans* (Giardina, 2006). *Micromonospora* sp. strain M42 ABC transporter sequences include 74 ABC permeases and 89 ATP-binding proteins. By comparison, 81 ABC permeases and 141 ATP-binding proteins were found in the genome of *S. coelicolor* A3(2) (Bentley *et al.*, 2002). From a nutritional point, sponges are thought to be suitable ecological niches for microbes when compared with the nutrient-poor seawater. This is especially true for tropical seawater and it has been proposed that because of the low level of nutrients available, mutualistic symbioses have developed (Muscatine & Porter, 1977). In sponges, ammonia is likely to be the main nitrogen source as it is an end product of the host metabolism, carbohydrates and amino acid should also be present because of the continual phagocytosis of the sponge (Hentschel *et al.*, 2006). In bacteria, ABC transporters are predominantly involved in nutrient uptake (Locher *et al.*, 2002) though they also participate in the export of toxins and harmful toxins (Nikaido, 1994).

The high number of ABC transporter sequences in the genome of *Micromonospora* sp. strain M42 indicates that this system is very active. Because of the known involvement

of ABC transporters in nutrient uptake, it is likely to perform a similar role for

Micromonospora sp. strain M42 and the abundance of ABC transporter sequences

suggests that *Micromonospora* sp. strain M42 can deal with a wide array of carbon

sources that may be available in the sponge mesohyl at any given time.

Drug resistance transporters were identified, including transporters of the Bcr/CflA

family that provides resistance to bicyclomycin, chloramphenicol and florfenicol, and the

DHA2 family, which provides resistance to tetracycline.

Among the identified transporters is a RhtB family transporter which can transport

homoserine, homoserine lactone and β-hydroxynorvaline and is likely to protect

Micromonospora sp. strain M42 against inhibitory concentrations of homoserine

(Livshits *et al.*, 2003). Homoserine was shown to inhibit nicotinamide adenine

dinucleotide phosphate (NADP$^+$)-specific glutamate dehydrogenase (EC 1.4.1.4), the

enzyme catalyzing the first reaction in ammonia assimilation (Kotre *et al.*, 1973). Both

homoserine and acyl homoserine lactone can inhibit microorganisms' growth (Kotre *et
al.*, 1973; Wilkinson *et al.*, 2002).

Acyl homoserine lactones or AHLs are used by many Gram-negative bacteria to

communicate (Camilli & Bassler, 2006) and have been found to be produced by sponge-

associated bacteria (Taylor *et al.*, 2004). The concentration of AHLs within the sponge is

unclear. The authors found that one out of the 11 Gram negative bacteria tested was an

AHL producer (Taylor *et al.*, 2004). We can assume that because of the difficulty to

isolate sponge-associated microbe that this ratio is an underestimate of the number of

bacterial AHLs producers within the sponge which could a high concentration of AHLs

within the sponge tissue.

Though some acyl homoserine lactones can diffuse freely through membranes, others require active transfer through efflux pumps (Pearson *et al.*, 1999). Several Gram-positive bacteria were found to produce AHL-inactivating enzymes (Dong *et al.*, 2002; Park *et al.*, 2006) and some Gram-positive bacteria such as *S. aureus* are sensitive to AHLs, influencing its growth and exotoxin production (Qazi *et al.*, 2006). No Gram-positive bacteria have been found to produce AHLs, yet they can be sensitive to AHLs and it is possible that RhtB type transporters are involved in the active transport of these quorum sensing molecules providing in the case of *Micromonospora* sp. strain M42 with information about its environment.

Other transporters for which the substrate is predictable include those for sugars, amino acids, peptides, metals and ions, indicating that M42 might maximize the nutrients available from the sponge.

5.3.3 Two-component regulatory system

The two-component response regulatory system is one of the main mechanisms used by bacteria to sense the changes occurring in their environment (Stock *et al.*, 2000). It is composed of a sensor kinase, usually located in the membrane and a response regulator (Galperin, 2004). The stimulus is detected by the sensor kinase, which is also responsible for the transduction signal via phosphorylation to the partner response regulator. The response regulator then formulates an appropriate response. By being constantly aware of their surroundings, bacteria can adapt by altering the expression of specific genes or modify other cellular activities. It has been shown that there is a

correlation between the number of two component response regulator genes and both the genome complexity and the lifestyle of the organism (Ashby, 2004). Bacteria living in complex habitats such as soil or aqueous environments have about twice as many response regulator genes than bacteria living in more stable habitats like symbionts and pathogens (Ashby, 2004). The genome size plays also an important role and bacteria with larger genomes have more response regulator genes (Ashby, 2004). *Micromonospora* sp. strain M42 possesses a high number of sequences of two-component regulator genes that includes 73 sensor kinases (histidine and serine/threonine kinases) and 75 response regulators. The genome had members of the previously described regulator families LuxR, GntR, AfsR, LacI, AraC, MerR, TetR, GlnR, ArsR and IclR. The habitat of *Micromonospora* sp. strain M42 is the sponge mesohyl. Sponges are more than any other invertebrates exposed to their environment as they are not only immersed in it but also filter very large volumes of seawater (up to 24,000 l of seawater per kg of sponge per day) (Vogel, 1977). It is unclear how stable and controlled the mesohyl environment is. However, it can be hypothesized that because of the stability of the seawater itself, it should be relatively constant. This would likely not require sponge symbionts to have a high number of two-component regulatory genes. The fact that *Micromonospora* sp. strain M42 is a heterotrophic bacterium and that it has a rather large genome (6.7 Mb) could explain the numerous two-component regulatory genes present in its genome.

5.3.4 Amino acids biosynthesis

Analysis of *Micromonospora* sp. strain M42 genome revealed 182 sequences that were identified as playing a role in the biosynthesis of amino acids. Biosynthetic genes for 12 amino acids were present, including tryptophan, asparagine, lysine, methionine, threonine, glutamate, arginine, glutamine, proline, leucine, cysteine and serine. Though no complete amino acids biosynthesis pathways were found, it is reasonable to assume that *Micromonospora* sp. strain M42 is capable of fully synthesizing amino acids for which parts of the biosynthetic gene cluster were identified. A number of genes encoding for the shikimate pathway are present in the genome including *aroA*, *aroB*, *aroF*, *aroK* and *aroQ*. The shikimate pathway is used by bacteria to synthesize aromatic amino acids such as tryptophan, phenylalanine and tyrosine. It bridges metabolism of carbohydrates to biosynthesis of aromatic compounds. In seven sequential metabolic steps, phosphoenolpyruvate and erythrose 4-phosphate are converted to chorismate, the precursor of the aromatic amino acids and many aromatic secondary metabolites such as alakaloids. The shikimate pathway is found only in microorganisms and plants, never in animals (Herrmann & Weaver, 1999). One common precursor of alkaloids and nitrogenous microbial metabolites is *p*-aminobenzoic acid (PABA) (Figure 5.5). Its biosynthesis derives from chorismate in the shikimate pathway. A study on the aureothin gene cluster of *Stretomyces thioluteus* demonstrated that PABA is a precursor to the novel polyketide synthase starter unit p-nitrobenzoate (PNBA) (Robert Winkler, 2005) emphasizing further its importance in the synthesis of a variety of microbial bioactive compounds.

157
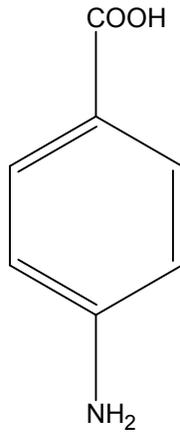
Figure 5.5. *p*-amino benzoic acid

Synthesis of tryptophan, phenylalanine and tyrosine begin with the synthesis of

chorismate, an important intermediate for many biosynthetic pathways (Figure 5.6).

Figure 5.6. Phenylalanine, tyrosine and tryptophan biosyntheses pathway

Phosphoenol pyruvate and erythrose 4-phosphate serve as beginning substrates for the

pathway. Trytophan synthesis is complex and involves 5 steps from chorismate (Figure

5.7). Glutamate donates an amine group in the first step of the pathway and pyruvate is

lost from chorismate. In the next threes steps a ribose sugar is added, this eventually

contributes to the 5 membered ring of tryptophan. Energy is contributed to the process in

the form of hydrolysis of pyrophosphate. This hydrolysis helps drive the addition of the

ribose sugar in the second step of the reaction. In the last step of the pathway serine

serves as the donor of the alpha carbon amino group of tryptophan.

Figure 5.7. Synthesis of tryptophan

Multiple sequences for tryptophan synthesis genes were found that could indicate multiple loci for tryptophan biosynthesis. Tryptophan is the amino acid that is a highly likely precursor for the synthesis of manzamine A (Baldwin & Whitehead, 1992). The analysis of the *S. coelicolor* genome showed that a tryptophan locus was included within the gene cluster for production of a calcium-dependent antibiotic (CDA) that contains tryptophan (Bentley *et al.*, 2002). It was proposed that the inclusion of a tryptophan locus within the CDA cluster would ensure an adequate supply of tryptophan, independently of the needs of protein synthesis. By analogy, it is possible that *Micromonospora* sp. strain M42 harbor a similar arrangement in the manzamine A biosynthesis pathway with the inclusion of a tryptophan biosynthetic locus.

5.3.5 Secondary metabolite capabilities

Actinomycetes are well known for their ability to produce secondary metabolites. A model analysis of the possible number of antibiotics that could be produced by the genus *Streptomyces* suggested it to be on the order of 100,000 (Watve *et al.*, 2001). The

percentage of the genome devoted to the production of secondary metabolites varies between organisms. *S. coelicolor* possesses 23 biosynthetic gene clusters predicted to encode for secondary metabolites, representing 4% of its total genome (Bentley *et al.*, 2002). *S. avermitilis* devotes an even larger part of its genome to the production of secondary metabolites with 30 gene clusters accounting for about 6% of its genome (Ikeda *et al.*, 2003). The analysis of *Micromonospora* sp. strain M42 genome showed that 269 sequences representing 3.6% of the total number of sequences were predicted to play a role in the production of secondary metabolites. Of these, 269 sequences, 45% (121) are polyketide synthase (PKS) sequences confirming the results obtained using a molecular approach described in Chapter 4 and 15% (40) are non-ribosomal peptide synthetases. Other polyketide-related sequences included acyltransferases that are part of the loading module along with acyl-carrier-proteins and are also found in the elongation module. Other related secondary metabolite sequences of interest present in *Micromonospora* sp. strain M42 genome include lantibiotic ABC transporter ATP binding/permease protein and LytB protein. Lantibiotics are heavily-modified ribosomally synthesized bacteriocin-like peptides produced by Gram-positive bacteria that can have their own dedicated transport system (Biet *et al.*, 1998). The Lyt B protein is involved in the non-mevalonate pathway of isoprenoid biosynthesis (Cunningham *et al.*, 2000). Isoprenoids represent the largest family of natural products with more than 22,000 structures known, many of them isolated from marine sources such as corals, algae and sponges (Kashman & Rudi, 2004).

This abundance and diversity of secondary metabolite genes in *Micromonospora* sp. strain M42 genome is not surprising for a member of the actinomycetes and demonstrates

the potential of the bacterium to produce a variety of natural products. There is strong

evidence that *Micromonospora* sp. strain M42 is capable of producing polyketide and

NRPS compounds. The pathways for these compounds could be fully investigated by

restriction mapping and sequencing selected fosmid clones containing these genes.

Furthermore, the structure of the metabolites could be predicted based on the sequence of

the pathway and the organization of the modules as it was done with the prediction of the

siderophore coelichelin based on *S. coelicolor* genome (Challis & Ravel, 2000) or the

discovery of the anti-fungal ECO-02301 from *Streptomyces aizunensis* (McAlpine *et al.*,

2005). It is possible that the secondary metabolite pathways may be truncated, inactive

or dormant. Under the right growth conditions, it may be possible to re-activate some of

them and observe their production by *Micromonospora* sp. strain M42. Another

approach would be to clone these pathways in a heterologous host which would offer

better control over the production of the compounds and likely better yields. For some of

the pathways encoding for polyketides or non-ribosomal peptides, their cloning into a

heterologous host can be challenging because of their large size, e.g. the CDA pathway in

S. coelicolor is 80 kb. Pathways of this size are likely to be spread over several fosmids.

However, reconstruction of large pathways in BAC is possible (Sosio *et al.*, 2001)

providing a solution for the expression of large secondary metabolite pathways.

Overall, while it was not possible to identify the manzamine biosynthetic pathway in the

limited sequence dataset, the analysis revealed the enormous potential of

Micromonospora sp. strain M42 to produce novel secondary metabolites.

5.3.6 Mobile elements

The investigation of *Micromonospora* sp. strain M42 genome revealed the presence of 39 mobile element sequences. The diversity of bacterial genomes is increased by the acquisition of new functions provided by genes that have been laterally transferred (Dutta & Pan, 2002). Genes acquired from lateral transfer provide a selective benefit to the recipient organism i.e. multiple antibiotic resistance or virulence (Daubin *et al.*, 2003). Genetic material can be acquired by transformation from naked DNA, by conjugation from a plasmid or by transduction, mediated by a bacteriophage. Sequences related to two types of bacteriophages, mycobacteriophage L5 and bacteriophage ΦC31 were identified in *Micromonospora* sp. strain M42 genome. These sequences have a high similarity with the phage sequences with E-value of 2.16E-30 for mycobacteriophage L5 and 1.86E-60 for bacteriophage ΦC31. For mycobacteriophage L5, a sequence identified as *mIHF* gene was found. The *mIHF* gene encodes for the mycobacterial integration host factor or mIHF, and is required for the site-specific integration of mycobacteriophage L5 (Lee & Hatfull, 1993). The protein mIHF does not bind to L5 *att*P DNA but is required for the formation of recombinogenic intrasomes that contains *att*P DNA, L5 integrase and mIHF (Pedulla & Hatfull, 1998). Bacteriophage ΦC31 is a common temperate *Streptomyces* phage that has been at the origin of many tools for genetic engineering of *Streptomyces* including phage cloning vectors (Bruton *et al.*, 1991), site-specific cloning vectors (Kuhstoss *et al.*, 1991) and a cosmid cloning vector (Kobler *et al.*, 1991). The presence of bacteriophages ΦC31 and L5 related sequences indicates that *Micromonospora* sp. strain M42 may possess *att* sites that are required for the integration

of these bacteriophages into the chromosome. This discovery opens a wide range of possibilities to work on *Micromonospora* sp. strain M42 genome. The presence of *att* sites in *Micromonospora* sp. strain M42 genome would permit to use the well developed engineering tools for *Streptomyces* which would offer an alternative to random mutagenesis by making possible rational directed strain improvements possible. Also, the presence of phage sequences certainly indicates that the genome of *Micromonospora* sp. strain M42 is still evolving and continues to acquire genomic materials through phage infections.

The low-coverage genome sequence analysis of *Micromonospora* sp. strain M42 is an important step toward better comprehending how this bacterium functions. It provided much information about the microbe's genomic arrangement and metabolic capabilities. The analysis of the sequences did not provide any indication as to what genes may be involved in the biosynthesis of manzamines and other approaches such as transposon mutagenesis will be needed in order to identify them. This work confirmed that *Micromonospora* sp. strain M42 possesses polyketide genes as shown in Chapter 4 but also revealed many more sequences related to polyketide as well as many unknown non-ribosomal peptide genes. The fact that *Micromonospora* sp. strain M42 has the potential to produce additional natural products other than manzamines warrants further work aimed at identifying these additional compounds and their pathways. *Micromonospora* sp. strain M42 fosmid library can be used to fully sequence specific fosmid clones of interests such as those containing PKS or NRPS sequences. With a prominent transporter system and two-component system, *Micromonospora* sp. strain M42 seems to be well adapted to the sponge environment. *Micromonospora* sp.

strain M42 appears as a versatile microbe that can utilize a variety of nutrients including sugars, amino acids, peptides and also metals and ions that may be available in the sponge mesohyl. *Micromonospora* sp. strain M42 large number two-component system genes certainly helps the microbe to sense the variation of its immediate environment, probably mediating rapid switches to use new nutritional sources that ensures the strain thrives in a competitive environment such as the sponge mesohyl.

6 General discussion and future directions

Sponges are the source of more natural products and bioactive compounds than any other marine invertebrate (Osinga *et al.*, 1998). The similarity between some sponge compounds and metabolites known to be of microbial origin led to speculation that sponge-associated microbes, rather than the sponge itself, may be responsible for their production. Two sponges, *Acanthostrongylophora* Sponge 35 and Sponge 52 known to contain manzamine alkaloids were collected off the coast of Manado in Indonesia. Manzamine A is active against various infectious diseases including tuberculosis and malaria and was proposed to be of microbial origin (Kobayashi & Ishibashi, 1993). The aim of this study was to investigate the microbial community associated with the sponges and isolate a manzamine microbial producer.

The microbial communities of the Indonesian sponges were investigated using 16S rRNA gene (**Chapter 2**). The microbial community analysis of both sponges showed that they were both complex and diverse. The two sponges had in common several major bacterial groups including *Actinomycetes, Acidobacteria, Chloroflexi,* alpha- and gamma-proteobacteria. *Acanthostrongylophora* sp. Sponge 35 microbial community is dominated by *Acidobacteria* and *Acanthostrongylophora* sp.Sponge 52 is dominated by gamma-proteobacteria. Actinomycetes are strongly represented in both microbial communities demonstrating that sponges can be a reservoir for actinomycetes which is especially interesting considering the track record of natural products from these bacteria. The amount of clones closely related to previously identify sponge-associated bacteria

was 77% for *Acanthostrongylophora* Sponge 35 and 66% for *Acanthostrongylophora* sp.Sponge 52. These results are consistent with findings by Hentshel *et al* (2002) that there are sponge-specific microbial communities. Fluorescent *in situ* hybridization of *Acanthostrongylophora* sp. Sponge 35 tissue with labeled probes showed the abundance of sponge-associated bacteria in the mesohyl and confirmed the presence of *Actinobacteria* and gamma-proteobacteria. The results obtained from the 16S rRNA gene clone library permitted selection of appropriate media with an emphasis being made on the isolation of actinomycetes (**Chapter 3**). Sixteen heterotrophic bacteria were isolated from *Acanthostrongylophora*, including alpha-proteobacteria, firmicutes and actinomycetes. Interestingly, one of the alpha-proteobacteria, M31, is closely related to NW001 (Webster & Hill, 2001), the alpha-proteobacterium that dominates the culturable community of *R. odorabile*. NW001-like bacteria have been recovered from seven different genera of sponges from geographically distant locations (Enticknap *et al.*, 2006). Additionally, it has been reported that a NW001-like bacterium was transferred vertically between sponge generations through the larvae (Enticknap *et al.*, 2006). It was proposed that the function of NW001-like bacteria may be related to nutrient acquisition because of its spatial location near the choanocyte chambers (Webster & Hill, 2001). NW001-like bacteria seem to play a role important enough for the sponge to justify their ubiquitous presence in different sponge microbial communities as well as a direct transfer through the larvae.

All isolates recovered from *Acanthostrongylophora* were tested for manzamine production using LC-MS and H^{1}NMR. One of the isolates, identified by 16S rRNA gene analysis as *Micromonospora* sp. strain M42 which produces manzamine A and 8-hydroxy

manzamine, the compounds first identified in the sponge. This is the first report of an isolated sponge-associated bacterium that produces a metabolite with biomedical potential originally found in the sponge. Until this study, previous reports of sponge-associated bacteria implicated in the synthesis of compounds were supported by circumstantial evidence such as for the filamentous cyanobacterial *O. spongeliae* associated with the sponge *Lamellodysidea herbacea* and colocalized with sponge bioactive compounds (Unson & Faulkner, 1993; Unson *et al.*, 1994).

Micromonospora sp. strain M42 was further investigated using microbiological and molecular techniques (**Chapter 4**). A time course experiment to monitor the production of manzamine A by *Micromonospora* sp. strain M42 showed that manzamine A was detected as early as day 1 and peaked in the middle of the exponential growth phase. This result demonstrates that manzamine A is produced constitutively by *Micromonospora* sp. strain M42 unlike secondary metabolites that are synthesized during the plateau growth phase. Manzamine A has cytoxicity and anti-microbial properties yet at the level produced by *Micromonospora* sp. strain M42, it is unclear what the role of manzamine A is for the microbe. The fact that the bacterium is continually producing manzamine A suggests that it is needed for its development and probably acts as a defense compound. A test for salt requirement and its impact on manzamine production revealed that *Micromonospora* sp. strain M42 reaches a higher cell density, more rapidly with little or no salt in the ISP2 medium. The production of manzamine A matches the cell density of the organism and manzamine A is more abundant in media with little or no salt added. These results show that *Micromonospora* sp. strain M42 is not a true marine

organism as it does not require NaCl to grow but is halotolerant. The fact that the manzamine concentration is higher in media with low or no NaCl can be explained by the higher cell density. However, the results of the ratio OD_{600} to manzamine A showed that higher NaCl concentrations in the media seem to stimulate a higher level of manzamine A production per cell.

The manzamine are an intriguing family of alkaloids because of their bioactivities and their complex structure. Baldwin *et al.* (1992) proposed a biogenetic pathway for manzamine A that involves two dihydropyridine rings with an alkyl residue and a tryptophan unit. Focusing on the decarboxylation of the tryptophan unit, *Micromonospora* sp. strain M42 was investigated for the presence of a putative tryptophan decarboxylase enzyme that would catalyze the reaction yielding tryptamine using both a molecular and biochemistry approach. Neither approach was successful in identifying a tryptophan decarboxylase gene or the enzyme itself. Because the manzamine biogenetic pathway is speculative, there is no certitude that a tryptophan unit is involved in the synthesis and in the case that it would, the designed primers may not have sufficient homology to *Micromonospora* sp. strain M42 tryptophan decarboxylase gene sequence to detect it, due to incorrect codon usage or a lack of sequence conservation with the sequences on which the primers were based.

The original *Micromonospora* sp. strain M42 produces 1 g/l of manzamine A. UV mutagenesis was used to improve the yield. Selected hyper-producing mutants showed in the best case a 3.5 fold increase in manzamine A production. Interestingly, one of the mutants, M42-2 produces manzamine A and B. Manzamine B was not found in the

sponge suggesting that the mutagenesis is responsible for being able to detect manzamine B. Manzamine B was proposed as a precursor to manzamine A. The mutagenesis may have disrupted the final step in the manzamine A biosynthetic pathway, shortening it and causing the accumulation of intermediate products such as manzamine B. The fact that M42-2 retained the ability to produce manzamine A can be explained by a multiple number of copies of the manzamine A biosynthetic pathway in the genome. One of the pathways may have been mutated hence yielding manzamine B while the others are still producing manzamine A. Another possibility is that the mutagenesis affected the regulation level of the enzymes responsible for processing manzamine B which would result in a lesser number of enzymes than originally required, creating a bottleneck situation that makes manzamine B detectable. Finally, another explanation could be that a minor part of the mutant M42-2 population may have reverted to a pre-mutation stage, producing manzamine A. Subsequent mutations of the first generation mutants are likely to further improve the yield of manzamine to a level that would be more suitable for industrial standards (10 g/l). Manzamines are very promising drug leads because of their bioactivities yet the toxicity of the molecules can be an issue and prevent advancement to the clinical trial stages. Mutagenesis can be a way to generate other manzamines and analogs that may have different activities and lower toxicity. It has the advantage that it may lead to entirely new manzamines and also provide producing strains that will greatly reduce the cost of production of the molecule.

The genome size of *Micromonospora* sp. strain M42 was estimated by pulsed field gel electrophoresis to be around 6.7 Mb which is fairly large compared with other bacteria

though smaller than the 9 Mb genome found in the *Streptomyces* genus (Bentley *et al.*, 2002; Ikeda *et al.*, 2003). Because actinomycetes are well known for producing secondary metabolites including PKS and NRPS, *Micromonospora* sp. strain M42 was investigated for the presence of PKS type I synthase using a molecular approach. The PCR amplification of two different polyketide synthase type I sequences confirmed the hypothesis that *Micromonospora* sp. strain M42 has the genetic potential for the production of polyketides.

A fosmid library of *Micromonospora* sp. strain M42 genome was constructed (Chapter 5) and each ends of the fosmids were sequenced giving 7,435 reads. The sequences were compared to those in the GenBank and TIGR databases. All the cellular functions are represented including energy metabolism, transport and binding protein, central intermediary metabolism and secondary metabolism. Many sequences are linked to the ABC transporter system indicating that *Micromonospora* sp. strain M42 is very active and possesses a versatile nutrient uptake system. There is evidence that *Micromonospora* sp. strain M42 is resistant to several antibiotics including bicyclomycin, chloramphenicol, florfenicol and tetracycline. Two-component regulatory system sequences are well represented in *Micromonospora* sp. strain M42 genome with 73 sensor kinases and 75 response regulators. This shows that the bacterium can sense its surroundings and can likely respond and adapt to the changes that may occur.

Micromonospora sp. strain M42 possesses 216 sequences related to secondary metabolite production including polyketide synthases, peptide synthetases and acyltransferases. Considering the estimated genome size of *Micromonospora* sp. strain M42, it was

suspected that it may contain a number of genes devoted to secondary metabolism. The abundance of sequences and their diversity showed that *Micromonospora* sp. strain M42 has the genetic potential for the production of polyketides and non-ribosomal peptides. Because of the partial sequencing of the fosmids and the usually large size of polyketide and non-ribosomal peptide pathways, no complete pathway could be obtained.

The analysis of the *Micromonospora* sp. strain M42 fosmid library also revealed a number of mobile elements in its genome. Sequences related to bacteriophages L5 and ΦC31 were identified. Mobile elements are of interests because they can be used as genetic engineering tools. Bacteriophage ΦC31 is a common *Streptomyces* phage that have been used to create a variety of *Streptomyces* vectors including site specific cloning vectors.

In this study we report the isolation of a sponge-associated bacterium that produces a bioactive compound found in sponges. This study represents only the first step in a more detailed investigation of *Micromonospora* sp. strain M42 properties and capabilities. Though many efforts were devoted in identifying the manzamine biosynthetic pathway, this pathway remains elusive. The low pass genome sequencing analysis of *Micromonospora* sp. strain M42 provided valuable information about its genome composition. Although no genes could be attributed to a potential manzamine A biosynthetic pathway, we learned that *Micromonospora* sp. strain M42 genes were homologous to *Streptomyces* genes and that bacteriophages common to both *Mycobacterium* and *Streptomyces* were present in the genome. We hypothesize that it can be possible to use the genetic tools developed for *Streptomyces* for the genetic

engineering but also for mutation of *Micromonospora* sp. strain M42. By using

Streptomyces transposons such as Tn 4560 (Ikeda *et al.*, 1993) to mutate

Micromonospora sp. strain M42, it would be possible to disrupt the manzamine A

production and trace back the transposon location on the chromosome, this way

identifying the inactivated genes and uncover the manzamine A gene cluster. Once the

entire manzamine A biosynthetic pathway is elucidated, it will be possible to subclone it

in a heterologous host such as *E. coli* where the generation time is shorter than that of

Micromonospora sp. strain M42 and where production can be more easily controled The

yield of manzamine A can be improved substantially by submiting the regulation of the

manzamine A pathway to a super-promoter (Ni *et al.*, 1995) that will drive higher

expression. Manzamine A is a promising drug lead for malaria yet toxicity remains a

concern. The elucidation of the manzamine A biosynthetic genes opens possibility to

engineer the pathway to obtain metabolites like manzamine B and ircinal. Furthermore,

ircinol and ircinal, part of the building blocks of manzamine A, can be starting points of

tens of thousands of manzamine-derived molecules that can be obtained by semi-

synthesis. Having a microbe that can produce exclusively ircinol/ircinal would greatly

enhance the chances of obtaining a drug lead with reduced toxicity that may successfully

pass through clinical trials.

The partial sequence analysis of *Micromonospora* sp. strain M42 fosmid library provided

a wealth of information about the strain, its metabolism and how it interacts with its

environment. The complete sequencing of *Micromonospora* sp. strain M42 would permit

a deeper understanding not only of its secondary metabolite production capabilities but

also of its ecology and association with the sponge. Sponge-associated bacteria may be acquired from the surrounding water or they may be transmitted vertically. The mechanism of recognition and acquisition of sponge symbionts remains unclear. It was proposed that bacteria that resist digestion are more likely to become symbionts (Wilkinson *et al.*, 1984). It remains to be determined if a molecular mechanism is involved in the recognition process as it is for plant symbionts such as *Mesorhizobium loti* which possesses island of symbiosis. The analysis of *Micromonospora* sp. strain M42 genome may help in identifying genes that may be dedicated in the establishment and continuation of symbiosis.

The work presented in this dissertation shows that it is possible to isolate sponge-associated bacteria that produce bioactive metabolites originally ascribed to the sponge. This research suggests that microbiological approaches can provide a reliable solution to overcome the supply limitations so often encountered with marine natural products. The isolation of *Micromonospora* sp. strain M42 greatly improves the chances that manzamines will become a drug for the treatment of malarial.

7 Appendix 1

Phenol/Chloroform extraction

1. Extract the sample with two volumes of phenol, centrifuge at 6,000 x g for 10 min at room temperature to separate phases.

2. Transfer the aqueous phase into a clean tube

3. Add two volumes of chloroform/isoamyl alcohol, centrifuge at 6,000 x g for 10 min at room temperature to separate phases.

4. Transfer the aqueous phase into a clean tube and repeat chloroform extraction if supernatant is still cloudy.

8 Appendix 2

All quantities are for one liter of medium.

ISP2 medium (YME medium):

Yeast extract	4 g
Malt extract	10 g
Dextrose	4 g
NaCl	20 g

Starch casein agar medium:

Soluble starch	10 g
Casein (dissolved in NaOH)	1 g
K_2HPO_4	0.5 g
Agar	20 g
NaCl	20 g

Seawater starch agar medium:

Yeast extract	10 g
Peptone	4 g
Instant Ocean	29 g
KBr (concentration in seawater 0.08 g/l)	0.16 g
1 M Tris buffer pH 8.0	10 ml
Agar	20 g
Distilled water	990 ml

R2A medium:

Yeast extract	0.5 g
Difco proteose peptone no. 3	0.5 g
Casamino acids	0.5 g
Glucose	0.5 g
Soluble starch	0.5 g
Sodium pyruvate	0.3 g
K_2HPO_4	0.3 g
$MgSO_4. 7H_2O$	0.05 g
Agar	15 g
NaCl	20 g

Final pH 7.2 ; adjust with crystaline K_2HPO_4 or KH_2PO_4 before adding agar.

Chitin agar medium:

$(NH_4)_2SO_4$	2 g
KH_2PO_4	0.7 g
Na_2HPO_4	1.1g
$MgSO_4. 7H_2O$	.2g
$FeSO_4$	trace
$MnSO_4$	trace
Chitin, precipitated	3.0g
Agar	15 g
NaCl	20 g

Actinomycete isolation medium (AIA):

Sodium caseinate	2 g
Asparagine	0.1 g
Na(C$_2$H$_5$COO)	4 g
K$_2$HPO$_4$	0.5 g
MgSO$_4$. 7H$_2$O	0.1 g
FeSO$_4$. H$_2$O	0.001 g
Bacto agar	15 g
NaCl	20 g

Final pH 8.1 ± 0.2 at 25°C

Marine agar 2216 medium

Peptone	5.0 g
Yeast Extract	1.0 g
FeC$_6$H$_5$O$_7$	0.1 g
NaCl	19.45 g
MgCl$_2$. 6H$_2$O	8.8 g
Na$_2$SO$_4$	3.24 g
CaCl$_2$	1.8 g
KCl	0.55 g
NaHCO$_3$	0.16 g
KBr	0.08 g
SrCl$_2$	34.0 mg

H_3BO_3	22.0 mg
$2Na_2O.SiO_2$	4.0 mg
NaF	2.4 mg
NH_4NO_3	1.6 mg
HNa_2O_4P	8.0 mg
Agar	15.0 g

9   Appendix 3

Tryptophan Decarboxylase purification protocol

Adapted from Nakazawa *et al.* (1981) Aromatic L-amino acid decarboxylase from *Micrococcus percitreus*, purification, crystalisation and properties. Agr. Biol. Chem. 45(11), 2543-2552.

All operations are carried out at 0-5ºC unless specified.

Potassium phosphate buffers, pH 7.0 contain 10^{-5} M pyridoxal phosphate, 10^{-2}M 2-mercaptoethanol and 10^{-4} EDTA

1. Protein extraction with Sigma B-plus cell kit

2. First protamine treatment:

Add 1/5 of 10% protamine sulfate solution pH 7.0 dropwise while stirring

3. Centrifuge at 16,000 g for 30 min

4. Ammonium sulfate fractionation:

Add solid ammonium sulfate to give a 30% saturation

5. Centrifuge at 16,000 g for 30 min to remove precipitated proteins

6. Add solid ammonium sulfate to the supernatant to give a 50% saturation

7. Allow to stand overnight

8. Centrifuge at 16,000 g for 30 min

9. Dissolve in 0.05 M potassium phosphate buffer pH 7.0

10. Dialyze the solution for 24h during which 3 changes of 0.05 M potassium phosphate buffer pH 7.0 will be made

11. Second protamine treatment:
Add 15.37% (vol/vol) of 10% protamine sulfate at pH 7.0 to the dialyzate and stir for 1h

12. Centrifuge at 16,000 g for 30 min to remove the precipitate formed

13. Dialyze the supernatant overnight against 20 litters of 0.05 M potassium phosphate buffer pH 7.0

14. First DEAE-Sephadex column chromatography:
Apply the enzyme solution on a DEAE–Sephadex A-50 column (13x20 cm) equilibrated with 0.05 M potassium phosphate buffer pH 7.0

15. Wash thoroughly the column with 0.05 M potassium phosphate buffer pH 7.0

16. Elute the enzyme with 0.1 M potassium phosphate buffer pH 7.0 supplemented with 0.2 M KCl

17. Elute the enzyme with 0.1 M potassium phosphate buffer pH 7.0 supplemented with 0.3 M KCl

18. Pool the active fractions

19. Concentrate the active fractions by adding ammonium sulfate to give a saturation of 55%

20. Dissolve the precipitate in 0.05 M potassium phosphate buffer pH 7.0

21. Dialyze overnight against 100 times volume 0.05 M potassium phosphate buffer pH 7.0

22. Second ammonium sulfate fractionation:
Add solid ammonium sulfate to the dialyzed enzyme solution (500 ml)

23. Collect by centrifugation the precipitate formed between 30 and 38% saturation (16,000 g for 30 min)

24. Resuspend in 85 ml of 0.05 M potassium phosphate buffer pH 7.0

25. Dialyze against 0.05 M potassium phosphate buffer pH 7.0

26. Second DEAE-Sephadex column chromatography:

Apply the dialyzed enzyme solution (92 ml) on a DEAE-Sephadex A-50 column (4.4x14 cm) equilibrated with 0.05 M potassium phosphate buffer pH 7.0

27. Wash the column with 0.05 M potassium phosphate buffer pH 7.0

28. Elute the enzyme with a linear gradient of KCl concentration from 0.1 to 0.3 M in 0.1 M potassium phosphate buffer pH 7.0

29. Combine and concentrate the active fractions by addition of ammonium sulfate (55% saturation)

30. Dissolve the precipitated enzyme in 0.05 M potassium phosphate buffer pH 7.0

31. First Sephadex G-200 column chromatography:

Pass the enzyme solution through a Sephadex G-200 column (2.4x90 cm) equilibrated with 0.05 M potassium phosphate buffer pH 7.0

32. Combine the active fractions and concentrate them with ammonium sulfate (55% saturation)

33. Dissolve the precipitated enzyme in 0.05 M potassium phosphate buffer pH 7.0

34. Second Sephadex G-200 column chromatography:

Pass the enzyme solution (0.5 ml) through a Sephadex G-200 column (2.4x90 cm) equilibrated with 0.05 M potassium phosphate buffer pH 7.0

35. Combine the active fractions and concentrate them with ammonium sulfate (55% saturation)

36. Dissolve the precipitated enzyme in 0.05 M potassium phosphate buffer pH 7.0, containing 10^{-3} M dithiothreitol

37. Third ammonium sulfate fractionation:

Precipitate the enzyme with ammonium sulfate (55% saturation) and collect it in a 1.5 ml tube by centrifugation (16,000 g for 30 min)

38. Pool the active fractions (extracted between 30 and 35% saturation) and concentrate them with ammonium sulfate (50% saturation)

39. Dissolve the purified enzyme in 1 ml of 0.1 M potassium phosphate buffer pH 7.0

40. Dialyze overnight against 500 volumes of 0.1 M potassium phosphate buffer pH 7.0, containing 10^{-4} M pyridoxal phosphate, 10^{-2} M 2-mercaptoethanol, 10^{-3} M dithiothreitol, 10^{-4} EDTA and 20% glycerol.

41. Store at 4ºC

10 Appendix 4

Large-scale isolation of actinomycete total DNA

Reference: Hopwood, D. A. , M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf. Genetic manipulation of streptomycetes: a laboratory manual. The John Innes Foundation, Norwich.

Cultures:

Inoculate spore suspension/cell material scraped off ISP2/SC plates into two flasks containing 200 ml of ISP2 or SC. Flasks contained a S/S spring. Grow at 30˚C with shaking at 150 rpm for 72 h and remove 40 ml for DNA extraction from ISP2 flask.

Equipment:

Centrifuge

Waterbath

UV spectrophotometer for quantitation of DNA

Materials:

TE buffer

10% SDS

Proteinase K (20 mg/ml stock, stored at –20˚C)

5 M NaCl

CTAB/NaCl solution (10% CTAB in 0.7 M NaCl; dissolve 4.1 g NaCl in 80 ml water, slowly add 10 g CTAB [hexadecyltrimethyl ammonium bromide] while heating and stirring. Adjust final volume to 100 ml).

Chloroform/isoamyl alcohol (24:1)

Isopropanol

70% Ethanol

Protocol:

1. Harvest cells from 40 ml culture in sterile 50 ml tube by centrifugation at 4,000 g for 10 min, discard supernatant.

2. Resuspend pellet in 5 ml of TE buffer, add 10 mg lysozyme (i.e. to 2 mg/ml) and swirl to dissolve. Incubate at 30°C for 30 min-1 h, checking every 15 min. On first sign of increased viscosity, move to next step (or after 1 h).

3. Add 1.2 ml 0.5 M EDTA (i.e. to 0.1 M), mix gently and add 0.13 ml of Proteinase K (i.e. to 0.2 mg/ml), mix gently and incubate at 30°C for 5 min.

4. Add 6.5 ml 10% SDS (i.e. to 1%), tilt immediately and incubate at 37°C for up to 2 h, checking for lysis.

5. Add 1.2 ml of 5 M NaCl, mix, and add 1.0 ml of CTAB/NaCl solution.

6. Incubate at 65˚C for 20 min.

7. Extract with 2X volume of phenol, centrifuge at 6,000 g for 10 min at room temperature to separate phases.

8. Extract with 2X volume of chloroform/isoamyl alcohol, centrifuge at 6,000 g for 10 min at room temperature to separate phases.

9. Transfer supernatant to a clean tube and repeat chloroform extraction if supernatant is cloudy.

10. Precipitate DNA by addition of 0.6 to 1 volume of isopropanol and mix until a stringy white precipitate is visible. If possible, spool DNA onto a sealed glass Pasteur pipette, rinse in 70% ethanol, and proceed to step X.

11. Spin supernatant at 10,000 g for 15 min at 4˚C, discard supernatant.

12. Resuspend pellet in 70% ethanol to wash, spin at 10,000 g for 5 min. Discard supernatant and dry pellet in speedivac.

13. Resuspend pellet in 200 µl TE buffer (preferably overnight at 4˚C).

11 Appendix 5

Example of the analysis of sequences from *Micromonospora* sp. strain M42 fosmid library.

Read name	Read length	ORF_hit	Organism Hit	Common name	Gene symbol	EC number	Role id	E-value	Number of hit
FMMS625JF	1069	DR0921	Deinococcus radiodurans R1	trans-sulfuration enzyme			69	9.44E-36	1
FMMSB72JR	990	GHAORFA0013	Haloferax volcanii DS2	N-methylhydantoinase A	agaF	3.5.2.14	69	9.70E-48	1
FMMSF73JF	915	GHAORFA0014	Haloferax volcanii DS2	N-methylhydantoinase B (ATP-hydrolyzing)	huyB-1	3.5.2.14	69	2.54E-23	1
FMMSY21JR	918	GHAORFA0014	Haloferax volcanii DS2	N-methylhydantoinase B (ATP-hydrolyzing)	huyB-1	3.5.2.14	69	1.19E-28	1
FMMS280JR	1019	GSCORF1298	Streptomyces clavuligerus sequencing	anthranilate synthase		4.1.3.27	70	7.46E-35	1
FMMS393JF	1101	GMAORF3693	Mycobacterium avium 104	Shikimate kinase (SK)		2.7.1.71	70	1.03E-08	1
FMMS491JR	960	GANORF1245	Actinomyces naeslundii MG1	luxR family two component system response regulator			70	2.36E-35	1
FMMSA71JR	1043	GSCORF1298	Streptomyces clavuligerus sequencing	anthranilate synthase		4.1.3.27	70	5.28E-81	2
FMMSK89JF	1131	MSMEG2799	Mycobacterium smegmatis MC2	phospho-2-dehydro-3-deoxyheptonate aldolase		4.1.2.15	70	7.65E-72	2
FMMSV80JF	1055	GMAORF3693	Mycobacterium avium 104	Shikimate kinase (SK)		2.7.1.71	70	3.67E-16	4
FMMT190JR	1053	GMAORF3693	Mycobacterium avium 104	Shikimate kinase (SK)		2.7.1.71	70	7.88E-11	2
FMMT026JR	1062	GMAORF3693	Mycobacterium avium 104	Shikimate kinase (SK)		2.7.1.71	70	1.19E-14	6
FMMT458JF	1061	PFL4254	Pseudomonas fluorescens Pf-5	3-deoxy-7-phosphoheptulonate synthase		2.5.1.54	70	4.23E-12	1
FMMTB79JR	1079	GMAORF3693	Mycobacterium avium 104	Shikimate kinase (SK)		2.7.1.71	70	1.13E-18	6

References

Acinas, S. G., Klepac-Ceraj, V., Hunt, D. E., Pharino, C., Ceraj, I., Distel, D. L. & Polz, M. F. (2004). Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**, 551-554.

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V. & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl Environ Microbiol* **71**, 8966-8969.

Adrio, J. L. & Demain, A. L. (2006). Genetic improvement of processes yielding microbial products. *FEMS Microbiol Rev* **30**, 187-214.

Aizenberg, J., Weaver, J. C., Thanawala, M. S., Sundar, V. C., Morse, D. E. & Fratzl, P. (2005). Skeleton of *Euplectella* sp.: structural hierarchy from the nanoscale to the macroscale. *Science* **309**, 275-278.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

Andrade, A. C., Del Sorbo, G., Van Nistelrooy, J. G. M. & Waard, M. A. D. (2000). The ABC transporter AtrB from *Aspergillus nidulans* mediates resistance to all major classes of fungicides and some natural toxic compounds. *Microbiology* **146**, 1987-1997.

Andrianasolo, E. H., Gross, H., Goeger, D., Musafija-Girt, M., McPhail, K., Leal, R. M., Mooberry, S. L. & Gerwick, W. H. (2005). Isolation of swinholide A and related glycosylated derivatives from two field collections of marine cyanobacteria. *Org Lett* **7**, 1375-1378.

Ang, K. K., Holmes, M. J., Higa, T., Hamann, M. T. & Kara, U. A. (2000). In vivo antimalarial activity of the beta-carboline alkaloid manzamine A. *Antimicrob Agents Chemother* **44**, 1645-1649.

Arai, T., Yazawa, K. & Mikami, Y. (1976). Isolation and characterization of satellite antibiotics, mimosamycin and chlorocarcins from *Streptomyces lavendulae*, streptothricin source. *J Antibiot* **29**, 398-407.

Ashby, M. K. (2004). Survey of the number of two-component response regulator genes in the complete and annotated genome sequences of prokaryotes. *FEMS Microbiol Lett* **231**, 277-281.

Baldwin, J. E. & Whitehead, R. C. (1992). On the biosynthesis of manzamines. *Tetrahedron Lett* **33**, 2059-2062.

Bateman, A., Birney, E., Cerruti, L. & other authors (2002). The Pfam Protein Families Database. *Nucleic Acids Research* **30**, 276-280.

Becker, S., Boger, P., Oehlmann, R. & Ernst, A. (2000). PCR bias in ecological analysis: a case study for quantitative *Taq* nuclease assays in analyses of microbial communities. *Appl Environ Microbiol* **66**, 4945-4953.

Beja, O., Suzuki, M. T., Heidelberg, J. F., Nelson, W. C., Preston, C. M., Hamada, T., Eisen, J. A., Fraser, C. M. & DeLong, E. F. (2002). Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**, 630-633.

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M. & other authors (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.

Berdy, J. (1989). The discovery of new bioactive microbial metabolites: screening and identification. In *Bioactive microbial metabolites*, pp. 3-25. Edited by M. E. Bushell & U. Graefe. Amsterdam: Elsevier.

Bergquist, P. R. (1978). *Sponges*. Hutchinson, London.

Berquist, P. R. (1978). *Sponges*. London: University of California Press.

Berthold, R. J., Borowitzka, M. A. & Mackay, M. A. (1982). The ultrastructure of *Oscillatoria spongeliae*, the blue-green algal endosymbiont of the sponge *Dysidea herbacea*. *Phycologia* **21**, 327-335.

Bewley, C. A., Holland, N. D. & Faulkner, D. J. (1996). Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. *Experientia* **52**, 716-722.

Biet, F., Berjeaud, J. M., Worobo, R. W., Cenatiempo, Y. & Fremaux, C. (1998). Heterologous expression of the bacteriocin mesentericin Y105 using the dedicated transport system and the general secretion pathway. *Microbiology* **144**, 2845-2854.

Bjornsson, L., Hugenholtz, P., Tyson, G. W. & Blackall, L. L. (2002). Filamentous *Chloroflexi* (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. *Microbiology* **148**, 2309-2318.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A. & other authors (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1474.

Blunt, J. W., Copp, B. R., Munro, M. H., Northcote, P. T. & Prinsep, M. R. (2003). Marine natural products. *Nat Prod Rep* **20**, 1-48.

Blunt, J. W., Copp, B. R., Munro, M. H., Northcote, P. T. & Prinsep, M. R. (2004). Marine natural products. *Nat Prod Rep* **21**, 1-49.

Blunt, J. W., Copp, B. R., Munro, M. H., Northcote, P. T. & Prinsep, M. R. (2005). Marine natural products. *Nat Prod Rep* **22**, 15-61.

Borowitzka, M. A., Hinde, R. & Pironet, F. (1988). Carbon fixation by the sponge *Dysidea herbacea* and its endosymbiont *Oscillatoria spongeliae*. In *Proceedings of the 6th International Coral Reef Symposium*, pp. 151-155. Edited by J. H. Choat, D. J. Barnes, M. A. Borowitzka & other authors. Townsville, Australia: Symposium Executive Committee.

Boucher, Y., Nesbo, C. L. & Doolittle, W. F. (2001). Microbial genomes: dealing with diversity. *Curr Opin Microbiol* **4**, 285-289.

Bourguet-Kondracki, M. L., Martin, M. T. & Guyot, M. (1996). A new β-carboline alkaloid isolated from the marine sponge *Hyrtios erecta*. *Tetrahedron Lett* **37**, 3457-3460.

Brantley, S. E., Molinski, T. F., Preston, C. M. & DeLong, E. F. (1995). Brominated acetylenic fatty acids from *Xestospongia* sp., a marine spongebacteria association. *Tetrahedron* **51**, 7667-7672.

Breznak, J. A. (2002). Phylogenetic diversity and physiology of termite gut *Spirochetes*. *Integr Comp Biol* **42**, 313-318.

Bruton, C. J., Guthrie, E. P. & Chater, K. F. (1991). Phage vectors that allow monitoring of transcription of secondary metabolism genes in *Streptomyces*. *Biotechnology (N Y)* **9**, 652-656.

Bubb, M. R., Spector, I., Bershadsky, A. D. & Korn, E. D. (1995). Swinholide A is a microfilament disrupting marine toxin that stabilizes actin dimers and severs actin filaments. *J Biol Chem* **270**, 3463-3466.

Buchan, A., Gonzalez, J. M. & Moran, M. A. (2005). Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* **71**, 5665-5677.

Buchanan, R. E. (1917). Studies of the nomenclature and classification of the bacteria. II. The primary subdivisions of the Schizomycetes. *J Bacteriol* **2**, 155-164.

Buck, J. D. (1974). Effects of medium composition on the recovery of bacteria from sea water. In *Native aquatic bacteria: enumeration, activity and ecology*, pp. 19-28. Edited by J. W. Costerton & R. R. Colwell. Philadelphia: American Society for Testing Materials.

Bultel-Poncé, V., Debitus, C., Blond, A., Cerceau, C. & Guyot, M. (1997). Lutoside: an acyl-1-(acyl-6'-mannobiosyl)-3-glycerol isolated from the sponge-associated bacterium *Micrococcus luteus*. *Tetrahedron Lett* **38**, 5805-5808.

Bultel-Poncé, V., Berge, J.-P., Debitus, C., Nicolas, J.-L. & Guyot, M. (1999). Metabolites from the sponge-associated bacterium *Pseudomonas* species. *Mar Biotechnol* **1**, 384-390.

Bultel-Poncé, V. V., Debitus, C., Berge, J. P., Cerceau, C. & Guyot, M. (1998). Metabolites from the sponge-associated bacterium *Micrococcus luteus*. *J Mar Biotechnol* **6**, 233-236.

Burg, R. W., Miller, B. M., Baker, E. E. & other authors (1979). Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. *Antimicrob Agents Chemother* **15**, 361-367.

Camilli, A. & Bassler, B. L. (2006). Bacterial small-molecule signaling pathways. *Science* **311**, 1113-1116.

Casjens, S., Palmer, N., van Vugt, R. & other authors (2000). A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi. *Mol Microbiol* **35**, 490-516.

Chain, E. B. & Mellows, G. (1977). Pseudomonic acid. Part 1. The structure of pseudomonic acid A, a novel antibiotic produced by *Pseudomonas fluorescens*. *J Chem Soc [Perkin 1]*, 294-309.

Challis, G. L. & Ravel, J. (2000). Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol Lett* **187**, 111-114.

Chen, Y. S. (1960). Studies of the metabolic products of *Rosellinia necatrix* I. Isolation and characterization of several physiologically active neutral substances. *Bull Agr Chem Soc Japan* **24**, 372-381.

Chimeno, R. I. F., Canedo, L., Espliego, F., Gravalos, D. G., Calle, F., Puentes, J. L. F. & Romero, F. (2000). IB-96212, a novel cytotoxic macrolide produced by a marine *Micromonospora*. I. Taxonomy, fermentation, Isolation and biological activities. *J Antibiot* **53**, 474-478.

Chun, J. (1995). Computer-assisted classification and identification of actinomycetes. Ph.D. Thesis.

Cox, C. D. & Adams, P. (1985). Siderophore activity of pyoverdin for *Pseudomonas aeruginosa*. *Infect Immun* **48**, 130-138.

Crews, P., Cheng, X.-C., Adamczeski, M., Rodriguez, J., Jaspars, M., Schmitz, F. J., Traeger, S. C. & Pordesimo, E. O. (1994). 1,2,3,4-tetrahydro-8-hydroxymanzamines, alkaloids from two different haplosclerid sponges. *Tetrahedron* **50**, 13567-13574.

Cunningham, F. X., Jr., Lafond, T. P. & Gantt, E. (2000). Evidence of a role for LytB in the nonmevalonate pathway of isoprenoid biosynthesis. *J Bacteriol* **182**, 5841-5848.

Daubin, V., Lerat, E. & Perriere, G. (2003). The source of laterally transferred genes in bacterial genomes. *Genome Biol* **4**, R57.

Davidov, Y. & Jurkevitch, E. (2004). Diversity and evolution of *Bdellovibrio*-and-like organisms (BALOs), reclassification of *Bacteriovorax* starrii as *Peredibacter starrii* gen. nov., comb. nov., and description of the *Bacteriovorax-Peredibacter* clade as *Bacteriovoracaceae* fam. nov. *Int J Syst Evol Microbiol* **54**, 1439-1452.

Davidson, S. K., Allen, S. W., Lim, G. E., Anderson, C. M. & Haygood, M. G. (2001). Evidence for the biosynthesis of bryostatins by the bacterial symbiont "Candidatus Endobugula sertula" of the bryozoan Bugula neritina. *Appl Environ Microbiol* **67**, 4531-4537.

De Rosa, S., De Giulio, A., Tommonaro, G., Popov, S. & Kujumgiev, A. (2000). A beta-amino acid containing tripeptide from a *Pseudomonas-Alteromonas* bacterium associated with a Black Sea sponge. *J Nat Prod* **63**, 1454-1455.

Debitus, C., Guella, G., Mancini, I., Waikedre, J., Guemas, J. P., Nicolas, J. L. & Pietra, F. (1998). Quinolones from a bacterium and tyrosine metabolites from its host sponge, *Suberea creba* from the Coral Sea. *J Mar Biotechnol* **6**, 136-141.

Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* **27**, 4636-4641.

DeLong, E. F. (1992). Archaea in coastal marine environments. *Proc Natl Acad Sci U S A* **89**, 5685-5689.

Delong, E. F., Franks, D. G. & Yayanos, A. A. (1997). Evolutionary Relationships of Cultivated Psychrophilic and Barophilic Deep-Sea Bacteria. *Appl Environ Microbiol* **63**, 2105-2108.

DeLong, E. F., Preston, C. M., Mincer, T. & other authors (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**, 496-503.

Deng, W., Burland, V., Plunkett, G., 3rd & other authors (2002). Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* **184**, 4601-4611.

Devereux, R., He, S. H., Doyle, C. L., Orkland, S., Stahl, D. A., LeGall, J. & Whitman, W. B. (1990). Diversity and origin of *Desulfovibrio* species: phylogenetic definition of a family. *J Bacteriol* **172**, 3609-3619.

Donadio, S., Sosio, M. & Lancini, G. (2002). Impact of the first *Streptomyces* genome sequence on the discovery and production of bioactive substances. *Appl Microbiol Biotechnol* **60**, 377-380.

Dong, Y.-H., Gusti, A. R., Zhang, Q., Xu, J.-L. & Zhang, L.-H. (2002). Identification of quorum-quenching N-acyl homoserine lactonases from *Bacillus* species. *Appl Environ Microbiol* **68**, 1754-1759.

Duchaud, E., Rusniok, C., Frangeul, L. & other authors (2003). The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nat Biotechnol* **21**, 1307-1313.

Durand, P., Gros, O., Frenkiel, L. & Prieur, D. (1996). Phylogenetic characterization of sulfur-oxidizing bacterial endosymbionts in three tropical *Lucinidae* by using 16S rDNA sequence analysis. *Mol Mar Biol Biotechnol* **5**, 37-42.

Dutta, C. & Pan, A. (2002). Horizontal gene transfer and bacterial diversity. *J Biosci*, 27-33.

Edrada, R. A., Proksch, P., Wray, V., Witte, L., Muller, W. E. & Van Soest, R. W. (1996). Four new bioactive manzamine-type alkaloids from the Philippine marine sponge *Xestospongia ashmorica*. *J Nat Prod* **59**, 1056-1060.

Egan, S., Holmstrom, C. & Kjelleberg, S. (2001). *Pseudoalteromonas ulvae* sp. nov., a bacterium with antifouling activities isolated from the surface of a marine alga. *Int J Syst Evol Microbiol* **51**, 1499-1504.

El Sayed, K. A., Kelly, M., Kara, U. A., Ang, K. K., Katsuyama, I., Dunbar, D. C., Khan, A. A. & Hamann, M. T. (2001). New manzamine alkaloids with potent activity against infectious diseases. *J Am Chem Soc* **123**, 1804-1808.

Elyakov, G. B., Kuznetsova, T., Mikhailov, V. V., Maltsev, I. I., Voinov, V. G. & Fedoreyev, S. A. (1991). Brominated diphenyl ethers from a marine bacterium associated with the sponge *Dysidea* sp. *Experientia* **47**, 632-633.

Enticknap, J. J., Thompson, R., Peraud, O., Lohr, J. E., Hamann, M. T. & Hill, R. T. (2004). Molecular analysis of a Florida Keys sponge: Implications for natural products discovery. *Mar Biotechnol (NY)* **6**, S288-S293.

Enticknap, J. J., Kelly, M., Peraud, O. & Hill, R. T. (2006). Characterization of a culturable alphaproteobacterial symbiont common to many marine sponges and evidence for vertical transmission via sponge larvae. *Appl Environ Microbiol* **72**, 3724-3732.

Ermolyeva, Z. V., Ravich, I. V., Navashin, S. M. & other authors (1964). Antitumor activity of certain substances of natural origin. *Acta Unio Int Contr* **20**, 295-296.

Faulkner, D. J. (1986). Marine natural products. *Nat Prod Rep* **3**, 1-33.

Faulkner, D. J. (1987). Marine natural products. *Nat Prod Rep* **4**, 539-576.

Faulkner, D. J. (1988). Marine natural products. *Nat Prod Rep* **5**, 613-663.

Faulkner, D. J. (1990). Marine natural products. *Nat Prod Rep* **7**, 269-309.

Faulkner, D. J. (1991). Marine natural products. *Nat Prod Rep* **8**, 97-147.

Faulkner, D. J. (1993). Marine natural products. *Nat Prod Rep* **10**, 497-539.

Faulkner, D. J. (1994). Marine natural products. *Nat Prod Rep* **11**, 355-394.

Faulkner, D. J., M.D., U. & Bewley, C. A. (1994). The chemistry of some sponges and their symbionts. *Pure Appl Chem* **66**, 1983-1990.

Faulkner, D. J. (1996). Marine natural products. *Nat Prod Rep* **13**, 75-125.

Faulkner, D. J. (1998). Marine natural products. *Nat Prod Rep* **15**, 113-158.

Faulkner, D. J. (2000). Marine natural products. *Nat Prod Rep* **17**, 7-55.

Faulkner, D. J. (2001). Marine natural products. *Nat Prod Rep* **18**, 1-49.

Faulkner, D. J. (2002). Marine natural products. *Nat Prod Rep* **19**, 1-48.

Ferrer, J.-L., Jez, J. M., Bowman, M. E., Dixon, R. A. & Noel, J. P. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat Struct Mol Biol* **6**, 775-784.

Fleischmann, R. D., Adams, M. D., White, O. & other authors (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.

Flowers, A. E., Garson, M. J., Webb, R. I., Dumdei, E. J. & Charan, R. D. (1998). Cellular origin of chlorinated diketopiperazines in the dictyoceratid sponge *Dysidea herbacea* (Keller). *Cell Tissue Res* **292**, 597-607.

Fraser, C. M., Casjens, S., Huang, W. M. & other authors (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390**, 580-586.

Friedrich, A. B., Merkert, H., Fendert, T., Hacker, J., Proksch, P. & Hentschel, U. (1999). Microbial diversity in the marine sponge *Aplysina cavernicola* (formerly *Verongia cavernicola*) analyzed by fluorescence *in situ* hybridization (FISH). *Mar Biol* **134**, 461-470.

Fuerst, J. A., Webb, R. I., Garson, M. J., Hardy, L. & Reiswig, H. M. (1999). Membrane-bounded nuclear bodies in a diverse range of microbial symbionts of Great Barrier Reef sponges. *Mem Queensl Mus* **44**, 193-203.

Furstner, A., Grabowski, J., Lehmann, C. W., Kataoka, T. & Nagai, K. (2001). Synthesis and biological evaluation of nonylprodigiosin and macrocyclic prodigiosin analogues. *Chembiochem* **2**, 60-68.

Furumai, T., Takagi, K., Igarashi, Y., Saito, N. & Oki, T. (2000). Arisostatins A and B, new members of tetrocarcin class of antibiotics from *Micromonospora* sp. TP-A0316. I. Taxonomy, fermentation, isolation and biological properties. *J Antibiot (Tokyo)* **53**, 227-232.

Galperin, M. Y. (2004). Bacterial signal transduction network in a genomic perspective. *Environmental Microbiology* **6**, 552-567.

Gene Ontology, C. (2004). The Gene Ontology (GO) database and informatics resource. *Nucl Acids Res* **32**, D258-261.

Giardina, A. (2006). A putative membrane protein and an ABC-transporter of *Planobispora rosea* stimulate antibacterial activity in *Streptomyces lividans*. In *Engineering Conferences International*. Santa Fe New Mexico.

Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., Clardy, J., Goodman, R. M. & Handelsman, J. (2002). Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* **68**, 4301-4306.

Goericke, R. (2002). Bacteriochlorophyll *a* in the ocean: is anoxygenic bacterial photosynthesis important? *Limnol Oceanogr* **47**, 290-295.

Gokhale, R. S. & Tuteja, D. (2001). Biochemistry of polyketide synthases. *Biotechnology* **10**, 341-372.

Gotelli, N. J. & Entsminger, G. L. (2006). EcoSim: Null models software for ecology. Version 7. Edited by A. I. I. Kesey-Bear. Jericho, VT05465. http://garyentsminger.com/ecosim.htm.

Grant, R. E. (1836). *The encyclopedia of anatomy and physiology*. London: Sherwood, Gilbert and Piper.

Gunasekera, S. P., Gunasekera, M. & Longley, R. E. (1990). Discodermolide: a new bioactive polyhydroxylated lactone from the marine sponge *Discodermia dissoluta*. *J Org Chem* **55**, 4912-4915.

Guo, Y., Trivellone, E., Scognamiglio, G. & Cimino, G. (1998). Misenine, a novel macrocyclic alkaloid with an unusual skeleton from the Mediterranean sponge *Reniera* sp. *Tetrahedron* **54**, 541-550.

Haft, D. H., Selengut, J. D. & White, O. (2003). The TIGRFAMs database of protein families. *Nucl Acids Res* **31**, 371-373.

Hansen, G. H., Bergh, O., Michaelsen, J. & Knappskog, D. (1992). *Flexibacter ovolyticus* sp. nov., a pathogen of eggs and larvae of Atlantic halibut, *Hippoglossus hippoglossus L. Int J Syst Bacteriol* **42**, 451-458.

Hayashi, T., Makino, K., Ohnishi, M. & other authors (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**, 11-22.

Haygood, M. G., Schmidt, E. W., Davidson, S. K. & Faulkner, D. J. (1999). Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *J Mol Microbiol Biotechnol* **1**, 33-43.

He, H., Ding, W. D., Bernan, V. S., Richardson, A. D., Ireland, C. M., Greenstein, M., Ellestad, G. A. & Carter, G. T. (2001). Lomaiviticins A and B, potent antitumor antibiotics from *Micromonospora lomaivitiensis*. *J Am Chem Soc* **123**, 5362-5363.

He, J. & Hertweck, C. (2003). Iteration as programmed event during polyketide assembly; molecular analysis of the aureothin biosynthesis gene cluster. *Chem Biol* **10**, 1225-1232.

Heckly, R. J. (1964). Differentiation of exotoxin and other biologically active substances in *Pseudomonas pseudomallei* filtrates. *J Bacteriol* **88**, 1730-1736.

Heidelberg, J. F., Eisen, J. A., Nelson, W. C. & other authors (2000). DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. *Nature* **406**, 477-483.

Hentschel, U., Hopke, J., Horn, M., Friedrich, A. B., Wagner, M., Hacker, J. & Moore, B. S. (2002). Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl Environ Microbiol* **68**, 4431-4440.

Hentschel, U., Usher, K. M. & Taylor, M. W. (2006). Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* **55**, 167-177.

Hernandez, L. M. C., Blanco, J. A. F., Baz, J. P., Puentes, J. L. F., Millan, F. R., Vasquez, F. E., Chimeno, R. I. F. & Gravalos, D. G. (2000). 4'-N-Methyl-5'-

hydroxystaurosporine and 5'-hydroxystaurosporine, new indolocarbazole alkaloids from a marine *Micromonospora* sp. strain. *J Antibiot* **53**, 895-902.

Herrmann, K. M. & Weaver, L. M. (1999). The shikimate pathway. *Annu Rev Plant Physiol Plant Mol Biol* **50**, 473-503.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**, 67-113.

Hildebrand, M., Waggoner, L. E., Lim, G. E., Sharp, K. H., Ridley, C. P. & Haygood, M. G. (2004). Approaches to identify, clone, and express symbiont bioactive metabolite genes. *Nat Prod Rep* **21**, 122-142.

Hill, R. T. (2004). Microbes from marine sponges: a treasure trove of biodiversity for natural products discovery. In *Microbial Diversity and Bioprospecting*, pp. 177-190. Edited by A. T. Bull. Washington D.C: ASM press.

Hooper, J. N. A. & Van Soest, R. W. M. (2002). Class Demospongiae Sollas, 1885. In *Systema Porifera: a guide to the classification of sponges*, pp. 15-18. Edited by J. N. A. Hooper & R. W. M. Van Soest. New York: Kluwer Academic/Plenum

Hopwood, D. A. & Wright, H. M. (1983). CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **129**, 3575-3579.

Hopwood, D. A., Bibb, M. J., Chater, K. F. & other authors (1985). *Genetic manipulation of Streptomyces. A laboratory manual*. Norwich, U.K.: The John Innes Foundation.

Hu, J. F., Hamann, M. T., Hill, R. T. & Kelly, M. (2003). The manzamine alkaloids. In *The alkaloids*, pp. 207-285. Edited by G. A. Cordell: Elsevier.

Huber, T., Faulkner, G. & Hugenholtz, P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**, 2317-2319.

Hugenholtz, P. & Huber, T. (2003). Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases. *Int J Syst Evol Microbiol* **53**, 289-293.

Ichiba, T., Corgiat, J. M., Scheuer, P. J. & Kelly-Borges, M. (1994). 8-Hydroxymanzamine A, a beta-carboline alkaloid from a sponge, *Pachypellina* sp. *J Nat Prod* **57**, 168-170.

Iijima, M., Someno, T., Amemiya, M., Sawa, R., Naganawa, H., Ishizuka, M. & Takeuchi, T. (1999a). IC202A, a new siderophore with immunosuppressive activity produced by *Streptoalloteichus* sp. 1454-19. II. Physico-chemical properties and structure elucidation. *J Antibiot* **52**, 25-28.

Iijima, M., Someno, T., Imada, C., Okami, Y., Ishizuka, M. & Takeuchi, T. (1999b).
IC202A, a new siderophore with immunosuppressive activity produced by
Streptoalloteichus sp. 1454-19. I. Taxonomy, fermentation, isolation and biological
activity. *Journal of Antibiotics* **52**, 20-24.

**Iijima, M., Someno, T., Ishizuka, M., Sawa, R., Naganawa, H. & Takeuchi, T.
(1999c).** IC202B and C, new siderophores with immunosuppressive activity produced by
Streptoalloteichus sp. 1454-19. *J Antibiot* **52**, 775-780.

Ikeda, H., Takada, Y., Pang, C. H., Tanaka, H. & Omura, S. (1993). Transposon
mutagenesis by Tn4560 and applications with avermictin-producing *Streptomyces
avermitilis*. *J Bacteriol* **175**, 2077-2082.

**Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki,
Y., Hattori, M. & Omura, S. (2003).** Complete genome sequence and comparative
analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**,
526-531.

Imamura, N., Nishijima, M., Adachi, K. & Sano, H. (1993). Novel antimycin
antibiotics, urauchimycins A and B, produced by marine actinomycete. *J Antibiot* **46**,
241-246.

Inagaki, F., Nunoura, T., Nakagawa, S. & other authors (2006). Biogeographical
distribution and diversity of microbes in methane hydrate-bearing deep marine sediments
on the Pacific Ocean Margin. *Proc Natl Acad Sci U S A* **103**, 2815-2820.

Isnansetyo, A. & Kamei, Y. (2003). MC21-A, a bactericidal antibiotic produced by a
new marine bacterium, *Pseudoalteromonas phenolica* sp. nov. O-BC30T, against
methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* **47**, 480-488.

Ivanova, E. P. & Mikhailov, V. V. (2001). A new family of *Alteromonadaceae* fam.
nov., including the marine proteobacteria species *Alteromonas*, *Pseudoalteromonas*,
Idiomarina, and *Colwellia*. *Mikrobiologiia* **70**, 15-23.

**Ivanova, E. P., Sawabe, T., Lysenko, A. M., Gorshkova, N. M., Hayashi, K.,
Zhukova, N. V., Nicolau, D. V., Christen, R. & Mikhailov, V. V. (2002a).**
Pseudoalteromonas translucida sp. nov. and *Pseudoalteromonas paragorgicola* sp. nov.,
and emended description of the genus. *Int J Syst Evol Microbiol* **52**, 1759-1766.

**Ivanova, E. P., Shevchenko, L. S., Sawabe, T., Lysenko, A. M., Svetashev, V. I.,
Gorshkova, N. M., Satomi, M., Christen, R. & Mikhailov, V. V. (2002b).**
Pseudoalteromonas maricaloris sp. nov., isolated from an Australian sponge, and
reclassification of *Pseudoalteromonas aurantia* NCIMB 2033 as *Pseudoalteromonas
flavipulchra* sp. nov. *Int J Syst Evol Microbiol* **52**, 263-271.

Ivanova, E. P., Sawabe, T., Hayashi, K., Gorshkova, N. M., Zhukova, N. V., Nedashkovskaya, O. I., Mikhailov, V. V., Nicolau, D. V. & Christen, R. (2003). *Shewanella fidelis* sp. nov., isolated from sediments and sea water. *Int J Syst Evol Microbiol* **53**, 577-582.

Jayatilake, G. S., Thornton, M. P., Leonard, A. C., Grimwade, J. E. & Baker, B. J. (1996). Metabolites from an Antarctic sponge-associated bacterium, *Pseudomonas aeruginosa*. *J Nat Prod* **59**, 293-296.

Jeanthon, C. (2000). Molecular ecology of hydrothermal vent microbial communities. *Antonie Van Leeuwenhoek* **77**, 117-133.

Jensen, P. R., Dwight, R. & Fenical, W. (1991). Distribution of actinomycetes in near-shore tropical marine sediments. *Appl Environ Microbiol* **57**, 1102-1108.

Jensen, P. R. & Fenical, W. (1994). Strategies for the discovery of secondary metabolites from marine bacteria: ecological perspectives. *Annu Rev Microbiol* **48**, 559-584.

Jiang, Z., Boyd, K. G., Mearns-Spragg, A., Adams, D. R., Wright, P. C. & Burgess, J. G. (2000). Two diketopiperazines and one halogenated phenol from cultures of the marine bacterium *Pseudoalteromonas luteoviolacea*. *Nat Prod Lett* **14**, 435-440.

Kaeberlein, T., Lewis, K. & Epstein, S. S. (2002). Isolating "Uncultivable" Microorganisms in Pure Culture in a Simulated Natural Environment. *Science* **296**, 1127-1129.

Kanoh, K., Kamino, K., Leleo, G., Adachi, K. & Shizuri, Y. (2003). Pseudoalterobactin A and B, new siderophores excreted by marine bacterium *Pseudoalteromonas* sp. KP20-4. *J Antibiot* **56**, 871-875.

Kashman, Y. & Rudi, A. (2004). On the biogenesis of marine isoprenoids. *Phytochem Rev* **3**, 309-323.

Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C. & Richard Hutchinson, C. (1999). Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **284**, 1368-1372.

Kim, T. K., Garson, M. J. & Fuerst, J. A. (2005). Marine actinomycetes related to the 'Salinospora' group from the Great Barrier Reef sponge *Pseudoceratina clavata*. *Environ Microbiol* **7**, 509-518.

Kim, T. K., Hewavitharana, A. K., Shaw, P. N. & Fuerst, J. A. (2006). Discovery of a new source of rifamycin antibiotics in marine sponge actinobacteria by phylogenetic prediction. *Appl Environ Microbiol* **72**, 2118-2125.

Kitagawa, I., Kobayashi, M., Katori, T., Yamashita, M., Tanaka, J., Doi, M. & Ishida, T. (1990). Absolute stereostructure of swinholide A, a potent cytotoxic macrolide from the Okinawan marine sponge *Theonella swinhoei*. *J Am Chem Soc* **112**, 3710-3712.

Kluge, A. G. & Farris, F. S. (1969). Quantitative phyletics and the evolution of annurans. *Syst Zool* **18**, 1-32.

Kobayashi, J. & Ishibashi, M. (1993). Bioactive metabolites of symbiotic marine microorganisms. *Chem Rev* **93**, 8305-8308.

Kobler, L., Schwertfirm, G., Schmieger, H., Bolotin, A. & Sladkova, I. (1991). Construction and transduction of a shuttle vector bearing the cos site of *Streptomyces* phage phi C31 and determination of its cohesive ends. *FEMS Microbiol Lett* **62**, 347-353.

Kodaira, Y. (1961). Toxic substances to insects, produced by *Aspergillus ochraceus* and *Oospora destructor. Agric Biol Chem* **25**, 261-262.

Kolber, Z. S., Plumley, F. G., Lang, A. S. & other authors (2001). Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* **292**, 2492-2495.

Kon-ya, K., Shimidzu, N., Otaki, N., Yokoyama, A., Adachi, K. & Miki, W. (1995). Inhibitory effect of bacterial ubiquinones on the settling of barnacle, *Balanus amphitrite. Experientia* **51**, 153-155.

Koren-Goldshlager, G., Kashman, Y. & Schleyer, M. (1998). Haliclorensin, a novel diamino alkaloid from the marine sponge *Haliclonatulearensis. J Nat Prod* **61**, 282-284.

Kotre, A. M., Sullivan, S. J. & Savageau, M. A. (1973). Metabolic regulation by homoserine in *Escherichia coli* B/r. *J Bacteriol* **116**, 663-672.

Krinsky, N. I. (1994). The biological properties of carotenoids. *Pure Appl Chem* **66**, 1003-1010.

Kuhstoss, S., Richardson, M. A. & Rao, R. N. (1991). Plasmid cloning vectors that integrate site-specifically in *Streptomyces* spp. *Gene* **97**, 143-146.

Lafay, B., Ruimy, R., de Traubenberg, C. R., Breittmayer, V., Gauthier, M. J. & Christen, R. (1995). *Roseobacter algicola* sp. nov., a new marine bacterium isolated from the phycosphere of the toxin-producing dinoflagellate *Prorocentrum lima. Int J Syst Bacteriol* **45**, 290-296.

Lang, S., Beil, W., Tokuda, H., Wicke, C. & Lurtz, V. (2004). Improved production of bioactive glucosylmannosyl-glycerolipid by sponge-associated *Microbacterium* species. *Mar Biotechnol (NY)* **6**, 152-156.

Lee, E.-Y., Lee, H. K., Lee, Y. K., Sim, C. J. & Lee, J.-H. (2003). Diversity of symbiotic archaeal communities in marine sponges from Korea. *Biomolecular Engineering* **20**, 299-304.

Lee, H. K., Lee, D. S., Lim, J., Kim, J. S., Im, K. S. & Jung, J. H. (1998). Topoisomerase I inhibitors from the *Streptomyces* sp. strain KM86-9B isolated from a marine sponge. *Arch Pharm Res* **21**, 729-733.

Lee, M. H. & Hatfull, G. F. (1993). Mycobacteriophage L5 integrase-mediated site-specific integration in vitro. *J Bacteriol* **175**, 6836-6841.

Lee, Y. K., Lee, J.-H. & Lee, H. K. (2001). Microbial symbiosis in marine sponges. *J Microbiol* **39**, 254-264.

Li, C. W., Chen, J. Y. & Hua, T. E. (1998). Precambrian sponges with cellular structures. *Science* **279**, 879-882.

Liu, R., Cui, C. B., Duan, L., Gu, Q. Q. & Zhu, W. M. (2005). Potent in vitro anticancer activity of metacycloprodigiosin and undecylprodigiosin from a sponge-derived actinomycete *Saccharopolyspora* sp. nov. *Arch Pharm Res* **28**, 1341-1344.

Livshits, V. A., Zakataeva, N. P., Aleshin, V. V. & Vitushkina, M. V. (2003). Identification and characterization of the new gene rhtA involved in threonine and homoserine efflux in *Escherichia coli*. *Res Microbiol* **154**, 123-135.

Locher, K. P., Lee, A. T. & Rees, D. C. (2002). The *E. coli* BtuCD structure: A framework for ABC transporter architecture and mechanism. *Science* **296**, 1091-1098.

Lomovskaya, N., Otten, S. L., Doi-Katayama, Y. & other authors (1999). Doxorubicin overproduction in *Streptomyces peucetius*: cloning and characterization of the dnrU ketoreductase and dnrV genes and the doxA cytochrome P-450 hydroxylase gene. *J Bacteriol* **181**, 305-318.

Long, P. F., Dunlap, W. C., Battershill, C. N. & Jaspars, M. (2005). Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. *ChemBioChem* **6**, 1760-1765.

Lopez, J. V., McCarthy, P. J., Janda, K. E., Willoughby, R. & Pomponi, S. A. (1999). Molecular techniques reveal wide phyletic diversity of heterotrophic microbes associated with *Discodermia* spp. (Porifera: Demospongia). *Mem Queensl Mus* **44**, 329-341.

Lynn, T. & Callahan, I. (1974). Purification and characterization of *Pseudomonas aeruginosa* exotoxin. *Infect Immun* **9**, 113-118.

Maidak, B. L., Olsen, G. J., Larsen, N., Overbeek, R., McCaughey, M. J. & Woese, C. R. (1997). The RDP (Ribosomal Database Project). *Nucleic Acids Res* **25**, 109-111.

Maidak, B. L., Cole, J. R., Parker Jr, C. T. & other authors (1999). A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Res* **27**, 171-173.

Maldonado, L. A., Fenical, W., Jensen, P. R., Kauffman, C. A., Mincer, T. J., Ward, A. C., Bull, A. T. & Goodfellow, M. (2005). *Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family *Micromonosporaceae*. *Int J Syst Evol Microbiol* **55**, 1759-1766.

Manuel, M., Borojevic, R., Boury-Esnault, N. & Vacelet, J. (2002). Class Calcarea Bowerbank, 1864. In *Systema Porifera: a guide to the classification of sponges*, pp. 1103-1116. Edited by J. N. A. Hooper & R. W. M. Van Soest. New York: Kluwer Academic/Plenum

Matsuya, H., Okamoto, M., Ochi, T., Nishikawa, A., Shimizu, S., Kataoka, T., Nagai, K., Wasserman, H. H. & Ohkuma, S. (2000). Reversible and potent uncoupling of hog gastric (H(+)+K(+))-ATPase by prodigiosins. *Biochem Pharmacol* **60**, 1855-1863.

Matthijs, S., Baysse, C., Koedam, N. & other authors (2004). The *Pseudomonas* siderophore quinolobactin is synthesized from xanthurenic acid, an intermediate of the kynurenine pathway. *Mol Microbiol* **52**, 371-384.

McAlpine, J. B., Bachmann, B. O., Piraee, M., Tremblay, S., Alarco, A. M., Zazopoulos, E. & Farnet, C. M. (2005). Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. *J Nat Prod* **68**, 493-496.

McKee, T. C. & Ireland, C. M. (1987). Cytotoxic and antimicrobial alkaloids from the Fijian sponge *Xestospongia caycedoi*. *J Nat Prod* **50**, 754-756.

Meksuriyen, D. & Cordell, G. A. (1988). Biosynthesis of staurosporine, 1. 1H- and 13C-NMR assignments. *J Nat Prod* **51**, 884-892.

Meyerhans, A., Vartanian, J. P. & Wain-Hobson, S. (1990). DNA recombination during PCR. *Nucleic Acids Res* **18**, 1687-1691.

Miki, W., Yokoyama, A., Izumida, H. & Shimidzu, N. (1994). Okadaxanthin, a novel C_{50}-carotenoid from a bacterium, *Pseudomonas* sp. KK10206C associated with a marine sponge, *Halichondria okadai*. *Experientia* **50**, 684-686.

Miki, W., Otaki, N., Yokoyama, A. & Kusumi, T. (1996). Possible origin of zeaxanthin in the marine sponge, *Reniera japonica*. *Experientia* **52**, 93-96.

Miller, C. M., Miller, R. V., Garton-Kenny, D., Redgrave, B., Sears, J., Condron, M. M., Teplow, D. B. & Strobel, G. A. (1998). Ecomycins, unique antimycotics from *Pseudomonas viridiflava*. *J Appl Microbiol* **84**, 937-944.

Mincer, T. J., Jensen, P. R., Kauffman, C. A. & Fenical, W. (2002). Widespread and persistent populations of a major new marine actinomycete taxon in ocean sediments. *Appl Environ Microbiol* **68**, 5005-5011.

Mitova, M., Popov, S. & De Rosa, S. (2004a). Cyclic peptides from a *Ruegeria* strain of bacteria associated with the sponge *Suberites domuncula*. *J Nat Prod* **67**, 1178-1181.

Mitova, M., Tommonaro, G., Hentschel, U., Müller, W. E. G. & De Rosa, S. (2004b). Exocellular cyclic dipeptides from a *Ruegeria* strain associated with cell cultures of *Suberites domuncula*. *Marine Biotechnology* **6**, 95-103.

Montalvo, N. F., Mohamed, N. M., Enticknap, J. J. & Hill, R. T. (2005). Novel actinobacteria from marine sponges. *Antonie Van Leeuwenhoek* **87**, 29-36.

Montaner, B., Navarro, S., Pique, M., Vilaseca, M., Martinell, M., Giralt, E., Gil, J. & Perez-Tomas, R. (2000). Prodigiosin from the supernatant of *Serratia marcescens* induces apoptosis in haematopoietic cancer cell lines. *Br J Pharmacol* **131**, 585-593.

Moran, M., Rutherford, L. & Hodson, R. (1995). Evidence for indigenous *Streptomyces* populations in a marine environment determined with a 16S rRNA probe. *Appl Environ Microbiol* **61**, 3695-3700.

Morris, R. M., Rappe, M. S., Urbach, E., Connon, S. A. & Giovannoni, S. J. (2004). Prevalence of the *Chloroflexi*-related SAR202 bacterioplankton cluster throughout the mesopelagic zone and deep ocean. *Appl Environ Microbiol* **70**, 2836-2842.

Munro, M. H., Blunt, J. W., Dumdei, E. J., Hickford, S. J., Lill, R. E., Li, S., Battershill, C. N. & Duckworth, A. R. (1999). The discovery and development of marine compounds with pharmaceutical potential. *J Biotechnol* **70**, 15-25.

Muscatine, L. & Porter, J. W. (1977). Reef corals: Mutualistic symbioses adapted to nutrient-poor environment. *Bioscience* **27**, 454-460.

Muyzer, G., de Waal, E. C. & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**, 695-700.

Myers, E. W., Sutton, G. G., Delcher, A. L. & other authors (2000). A whole-genome assembly of Drosophila. *Science* **287**, 2196-2204.

Nakamura, H., Deng, S., Kobayashi, J. i., Ohizumi, Y., Tomotake, Y., Matsuzaki, T. & Hirata, Y. (1987). Keramamine-A and -B, novel antimicrobial alkaloids from the Okinawan marine sponge *Pellina* sp. *Tetrahedron Lett* **28**, 621-624.

Nelson, K. E., Clayton, R. A., Gill, S. R. & other authors (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323-329.

Newman, D. J. & Cragg, G. M. (2004). Marine natural products and related compounds in clinical and advanced preclinical trials. *J Nat Prod* **67**, 1216-1238.

Ni, M., Cui, D., Einstein, J., Narasimhulu, S., Vergara, C. E. & Gelvin, S. B. (1995). Strengh and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. *Plant J* **7**, 661-676.

Nikaido, H. (1994). Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* **264**, 382-388.

Nishina, A., Kihara, H., Uchibori, T. & Oi, T. (1991). Antimicrobial substances in "DF-100" extract of grapefruit seeds. *Bokin Bobai* **19**, 401-404.

Oclarit, J. M., Okada, H., Ohta, S., Kaminura, K., Yamaoka, Y., Iizuka, T., Miyashiro, S. & Ikegami, S. (1994). Anti-bacillus substance in the marine sponge, *Hyatella* species, produced by an associated *Vibrio* species bacterium. *Microbios* **78**, 7-16.

Okabe, S., Ito, T. & Satoh, H. (2003). Sulfate-reducing bacterial community structure and their contribution to carbon mineralization in a wastewater biofilm growing under microaerophilic conditions. *Appl Microbiol Biotechnol* **63**, 322-334.

Okami, Y. & Hotta, K. (1988). Search and discovery of new antibiotics. In *Actinomycetes in biotechnology*, pp. 33-67. Edited by M. Goodfellow, S. T. Williams & M. Mordarski. San Diego: Academic Press Inc.

Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., de Santos, V. & Cruz, L. J. (1985). Peptide neurotoxins from fish-hunting cone snails. *Science* **230**, 1338-1343.

Omura, S., Ikeda, H., Ishikawa, J. & other authors (2001). Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *Proc Natl Acad Sci U S A* **98**, 12215-12220.

Osinga, R., Tramper, J. & Wijffels, R. H. (1998). Cultivation of marine sponges for metabolite production: applications for biotechnology? *Trends Biotechnol* **16**, 130-134.

Park, S.-Y., Hwang, B.-J., Shin, M.-H., Kim, J.-A., Kim, H.-K. & Lee, J.-K. (2006). N-acylhomoserine lactonase producing *Rhodococcus* spp. with different AHL-degrading activities. *FEMS Microbiol Lett* **261**, 102-108.

Parkhill, J., Wren, B. W., Mungall, K. & other authors (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665-668.

Paul, V. J. (1992). *Ecological roles of marine natural products* Ithaca, New York: Cornell University Press.

Pearson, J. P., Van Delden, C. & Iglewski, B. H. (1999). Active efflux and diffusion are involved in transport of *Pseudomonas aeruginosa* cell-to-cell signals. *J Bacteriol* **181**, 1203-1210.

Pedulla, M. L. & Hatfull, G. F. (1998). Characterization of the mIHF gene of *Mycobacterium smegmatis*. *J Bacteriol* **180**, 5473-5477.

Peraud, O., Yousaf, M., Hamann, M. T. & Hill, R. T. (2002). Microbial community analysis of a marine sponge that contains the anti-malarial compound manzamine A. In *American Society for Microbiology 102nd General Meeting*. Salt Lake City, Utah.

Pettit, G. R., Herald, C. L., Doubeck, D. L. & Herald, D. L. (1982). Isolation and structure of bryostatin 1. *J Am Chem Soc* **104**, 6846-6848.

Petursdottir, S. K. & Kristjansson, J. K. (1997). *Silicibacter lacuscaerulensis* gen. nov., sp. nov., a mesophilic moderately halophilic bacterium characteristic of the Blue Lagoon geothermal lake in Iceland. *Extremophiles* **1**, 94-99.

Philp, R. B. (2001). Effects of experimental manipulation of pH and salinity on Cd^{2+} uptake by the sponge *Microciona prolifera* and on sponge cell aggregation induced by Ca^{2+} and Cd^{2+} *Arch Environ Contam Toxicol* **41**, 282-288.

Piel, J. (2002). A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc Natl Acad Sci USA* **99**, 14002-14007.

Piel, J. (2004). Metabolites from symbiotic bacteria. *Nat Prod Rep* **21**, 519-538.

Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N. & Matsunaga, S. (2004). Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci USA* **101**, 16222-16227.

Pitcher, D. G., Saunders, N. A. & Owen, R. J. (1989). Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* **8**, 151-156.

Preston, C. M., Wu, K. Y., Molinski, T. F. & DeLong, E. F. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov. *Proc Natl Acad Sci USA* **93**, 6241-6246.

Prokic, I., Brummer, F., Brigge, H., Gortz, H. D., Gerdts, G., Schutt, C., Elbrachter, M. & Muller, W. E. G. (1998). Bacteria of the genus *Roseobacter* associated with the toxic dinoflagellate *Procentrum lima*. *Protist* **149**, 347-357.

Proksch, P., Edrada, R. A. & Ebel, R. (2002). Drugs from the seas - current status and microbiological implications. *Appl Environ Microbiol* **59**, 125-134.

Proteau, P. J., Gerwick, W. H., Garcia-Pichel, F. & Castenholz, R. (1993). The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia* **49**, 825-829.

Qazi, S., Middleton, B., Muharram, S. H., Cockayne, A., Hill, P., O'Shea, P., Chhabra, S. R., Camara, M. & Williams, P. (2006). N-acylhomoserine lactones antagonize virulence gene expression and quorum sensing in *Staphylococcus aureus*. *Infect Immun* **74**, 910-919.

Rappe, M. S., Vergin, K. & Giovannoni, S. J. (2000). Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS Microbiol Ecol* **33**, 219-232.

Ravel, J., Schrempf, H. & Hill, R. T. (1998). Mercury resistance is encoded by transferable giant linear plasmids in two Chesapeake Bay *Streptomyces* strains. *Appl Environ Microbiol* **64**, 3383-3388.

Rawlings, B. J. (1997). Biosynthesis of polyketides. *Nat Prod Rep* **97**, 523-556.

Ray, T., Weaden, J. & Dyson, P. (1992). Tris-dependant site-specific cleavage of *Streptomyces lividans* DNA. *FEMS Microbiol Lett* **96**, 247-252.

Reiswig, H. M. (2002). Class Hexactinellida Schmidt, 1870. In *Systema Porifera: a guide to the classification of sponges*, pp. 1201-1210. Edited by J. N. A. Hooper & R. W. M. Van Soest. New York: Kluwer academic/Plenum

Reysenbach, A.-L., Giver, L. J., Wickham, G. S. & Pace, N. R. (1992). Differential amplification of rRNA genes by polymerase chain reaction. *Appl Environ Microbiol* **58**, 3417-3418.

Ridley, C. P., Bergquist, P. R., Harper, M. K., Faulkner, D. J., Hooper, J. N. & Haygood, M. G. (2005a). Speciation and biosynthetic variation in four dictyoceratid sponges and their cyanobacterial symbiont, *Oscillatoria spongeliae*. *Chem Biol* **12**, 397-406.

Ridley, C. P., John Faulkner, D. & Haygood, M. G. (2005b). Investigation of *Oscillatoria spongeliae* dominated bacterial communities in four Dictyoceratid sponges. *Appl Environ Microbiol* **71**, 7366-7375.

Riley, M. (1993). Functions of the gene products of *Escherichia coli*. *Microbiol Rev* **57**, 862-952.

Robert Winkler, C. H. (2005). Sequential enzymatic oxidation of aminoarenes to nitroarenes via hydroxylamines. *Angewandte Chemie International Edition* **44**, 4083-4087.

Romero, F., Espliego, F., Perez Baz, J., Garcia de Quesada, T., Gravalos, D., de la Calle, F. & Fernandez-Puentes, J. L. (1997). Thiocoraline, a new depsipeptide with antitumor activity produced by a marine *Micromonospora*. I. Taxonomy, fermentation, isolation, and biological activities. *J Antibiot (Tokyo)* **50**, 734-737.

Roosenberg, J. M., 2nd, Lin, Y. M., Lu, Y. & Miller, M. J. (2000). Studies and syntheses of siderophores, microbial iron chelators, and analogs as potential drug delivery agents. *Curr Med Chem* **7**, 159-197.

Rosamond, J. & Allsop, A. (2000). Harnessing the power of the genome in the search for new antibiotics. *Science* **287**, 1973-1976.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-425.

Sakai, R., Higa, T., Jefford, C. W. & Bernardinelli, G. (1986). Manzamine A; an antitumor alkaloid from a sponge. *J Am Chem Soc* **108**, 6404-6405.

Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* **94**, 441-448.

Sanger, F., Coulson, A. R., Friedmann, T. & other authors (1978). The nucleotide sequence of bacteriophage phiX174. *J Mol Biol* **125**, 225-246.

Santavy, D. L., Willenz, P. & Colwell, R. R. (1990). Phenotypic study of bacteria associated with the Caribbean sclerosponge, *Ceratoporella nicholsoni*. *Appl Environ Microbiol* **56**, 1750-1762.

Sato, T., Konno, H., Tanaka, Y., Kataoka, T., Nagai, K., Wasserman, H. H. & Ohkuma, S. (1998). Prodigiosins as a new group of H^+/Cl^- symporters that uncouple proton translocators. *J Biol Chem* **273**, 21455-21462.

Schirmer, A., Gadkari, R., Reeves, C. D., Ibrahim, F., DeLong, E. F. & Hutchinson, C. R. (2005). Metagenomic analysis reveals diverse polyketide synthase gene clusters in

microorganisms associated with the marine sponge *Discodermia dissoluta*. *Appl Environ Microbiol* **71**, 4840-4849.

Schmidt, E. W., Bewley, C. A. & Faulkner, D. J. (1998). Theopalauamide, a bicyclic glycopeptide from filamentous bacterial symbionts of the lithistid sponge *Theonella swinhoei* from Palau and Mozambique. *J Org Chem* **63**, 1254-1258.

Schmidt, E. W., Obraztsova, A. Y., Davidson, S. K., Faulkner, D. J. & Haygood, M. G. (2000). Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel delta-proteobacterium, "*Candidatus* Entotheonella palauensis". *Mar Biol* **136**, 969-977.

Schwarzer, D., Finking, R. & Marahiel, M. A. (2003). Nonribosomal peptides: from genes to products. *Nat Prod Rep* **20**, 275-287.

Shiba, T. (1991). *Roseobacter litoralis* gen. nov. sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll a. *Syst Appl Microbiol* **14**, 140-145.

Shick, J. M. & Dunlap, W. C. (2002). Mycosporine-like amino acids and related gadusols: biosynthesis, acumulation, and UV-protective functions in aquatic organisms. *Annu Rev Physiol* **64**, 223-262.

Shigemori, H., Bae, M., Yazawa, K., Sasaki, T. & Kobayashi, J. (1992). Alteramide A, a new tetracyclic alkaloid from a bacterium *Alteromonas* sp. associated with the marine sponge *Halichondria okadai*. *J Org Chem* **57**, 4317-4320.

Sigrist, C. J. A., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. & Bucher, P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief in Bioinformatics* **3**, 265-274.

Smith, W. (1875). *A Dictionary of Greek and Roman Antiquities* London: John Murray.

Somerville, C. C., Knight, I. T., Straube, W. L. & Colwell, R. R. (1989). Simple, rapid method for direct isolation of nucleic acids from aquatic environments. *Appl Environ Microbiol* **55**, 548-554.

Sosio, M., Bossi, E. & Donadio, S. (2001). Assembly of large genomic segments in artificial chromosomes by homologous recombination in *Escherichia coli*. *Nucleic Acids Res* **29**, E37.

Stackebrandt, E., Rainey, F. A. & Ward-Rainey, N. L. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* **47**, 479-491.

Staley, J. T. & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* **39**, 321-346.

Stevenson, C. S., Capper, E. A., Roshak, A. K. & other authors (2002). The identification and characterization of the marine natural product scytonemin as a novel antiproliferative pharmacophore. *J Pharmacol Exp Ther* **303**, 858-866.

Stierle, A. C., Cardellina, J. H. I. & Singleton, F. L. (1988). A marine *Micrococcus* produces metabolites ascribed to the sponge *Tedania ignis*. *Experientia* **44**, 1021.

Stierle, D. B. & Stierle, A. A. (1992). Pseudomonic acid derivatives from a marine bacterium. *Experientia* **44**, 1165-1169.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* **69**, 183-215.

Sullivan, J. T., Trzebiatowski, J. R., Cruickshank, R. W. & other authors (2002). Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J Bacteriol* **184**, 3086-3095.

Sundar, V. C., Yablon, A. D., Grazul, J. L., Ilan, M. & Aizenberg, J. (2003). Fibre-optical features of a glass sponge. *Nature* **424**, 899-900.

Sutton, G., White, O., Adams, M. & Kerlavage, A. R. (1995). TIGR Assembler: A new tool for assembling large shotgun sequencing projects. *Genome Science and Technology* **1**, 9-19.

Suzuki, M. T. & Giovannoni, S. J. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol* **62**, 625-630.

Syed, I. B. (2002). Islamic medicine: 1000 years ahead of its times. *J Islamic Med Asso* **2**, 2-9.

Tatusova, T. A. & Madden, T. L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett* **174**, 247-250.

Taylor, M. W., Schupp, P. J., Baillie, H. J., Charlton, T. S., de Nys, R., Kjelleberg, S. & Steinberg, P. D. (2004). Evidence for acyl homoserine lactone signal production in bacteria associated with marine sponges. *Appl Environ Microbiol* **70**, 4387-4389.

ter Haar, E., Kowalski, R. J., Hamel, E., Lin, C. M., Longley, R. E., Gunasekera, S. P., Rosenkranz, H. S. & Day, B. W. (1996). Discodermolide, a cytotoxic marine agent that stabilizes microtubules more potently than taxol. *Biochemistry* **35**, 243-250.

Tsuji, R. F., Yamamoto, M., Nakamura, A., Kataoka, T., Magae, J., Nagai, K. & Yamasaki, M. (1990). Selective immunosuppression of prodigiosin 25-C and FK506 in the murine immune system. *J Antibiot* **43**, 1293-1301.

Tyson, G. W., Chapman, J., Hugenholtz, P. & other authors (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37-43.

Unson, M. D. & Faulkner, D. J. (1993). Cyanobacterial symbiont synthesis of chlorinated metabolites from *Dysidea herbacea* (Porifera). *Experientia* **44**, 1021-1022.

Unson, M. D., Holland, N. D. & Faulkner, D. J. (1994). A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. *Mar Biol* **119**, 1-11.

Vacelet, J. (1970). Description de cellules à bacteries intranucléaires chez des éponges *Verongia*. *J Microsc (Paris)* **9**, 333-346.

Venkateswaran, K. & Dohmoto, N. (2000). *Pseudoalteromonas peptidolytica* sp. nov., a novel marine mussel-thread-degrading bacterium isolated from the Sea of Japan. *Int J Syst Evol Microbiol* **50**, 565-574.

Venter, J. C., Remington, K., Heidelberg, J. F. & other authors (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66-74.

Vogel, S. (1977). Current-induced flow through living sponges in nature. *Proc Natl Acad Sci USA* **74**, 2069-2071.

Von Wintzingerode, F., Gobel, U. B. & Stackebrandt, E. (1997). Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol Rev* **21**, 213-229.

Waksman, S. A. & Curtis, S. A. (1916). The actinomycetes of the soil. *Soil Sci* **1**, 99-134.

Waksman, S. A. & Curtis, S. A. (1918). The occurence of actinomycetes in the soil. *Soil Sci* **6**, 309-319.

Waksman, S. A. & Woodruff, H. B. (1940). The soil as a source of microorganisms antagonistic to disease-producing bacteria. *J Bacteriol* **40**, 581-600.

Watve, M. G., Tickoo, R., Jog, M. M. & Bhole, B. D. (2001). How many antibiotics are produced by the genus *Streptomyces*? *Arch Microbiol* **176**, 386-390.

Webster, N., Watts, J. & Hill, R. (2001a). Detection and phylogenetic analysis of novel crenarcheote and euryarcheote 16S rRNA gene sequences from a Great Barrier Reef sponge. *Mar Biotechnol.*

Webster, N. S. & Hill, R. T. (2001). The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an alpha-proteobacterium. *Mar Biol* **138**, 843-851.

Webster, N. S., Wilson, K. J., Blackall, L. L. & Hill, R. T. (2001b). Phylogenetic diversity of bacteria associated with the marine sponge *Rhopaloeides odorabile. Appl Environ Microbiol* **67**, 434-444.

Webster, N. S., Negri, A. P., Webb, R. I. & Hill, R. T. (2002). A spongin-boring α-proteobacterium is the etiological agent of disease in the Great Barrier Reef sponge *Rhopaloeides odorabile. Mar Ecol Prog Ser* **232**, 305-309.

Weinstein, M. J., Luedemann, G. M., Oden, E. M. & other authors (1963). Gentamicin, a new antibiotic complex from *Micromonospora. J Med Chem* **6**, 463-464.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**, 697-703.

Weyland, H. (1969). Actinomycetes in North Sea and Atlantic ocean sediments. *Nature (London)* **223**, 858.

Weyland, H. (1981). Characteristics of actinomycetes isolated from marine sediments. *Zentralblatt fur Bakteriologie Supplement* **11**, 309-314.

Wicke, C., Huners, M., Wray, V., Nimtz, M., Bilitewski, U. & Lang, S. (2000). Production and structure elucidation of glycoglycerolipids from a marine sponge-associated *Microbacterium* species. *J Nat Prod* **63**, 621-626.

Wilkinson, A., Danino, V., Wisniewski-Dye, F., Lithgow, J. K. & Downie, J. A. (2002). N-Acyl-Homoserine Lactone Inhibition of Rhizobial Growth Is Mediated by Two Quorum-Sensing Genes That Regulate Plasmid Transfer. *J Bacteriol* **184**, 4510-4519.

Wilkinson, C. R. (1978a). Microbial associations in sponges. III. Ultrastructure of the in situ associations of coral reef sponges. *Mar Biol* **49**, 177-185.

Wilkinson, C. R. (1978b). Microbial associations in sponges. II. Numerical analysis of sponge and water bacterial populations. *Mar Biol* **49**, 169-176.

Wilkinson, C. R. (1978c). Microbial associations in sponges. I. Ecology, physiology and microbial populations of coral reef sponges. *Mar Biol* **49**, 161-167.

Wilkinson, C. R. & Garrone, R. (1980). Nutrition of marine sponges. Involvement of symbiotic bacteria in the uptake of dissolved carbon. In *Nutrition in the lower metazoa*, pp. 157-161. Edited by D. C. Smith & Y. Tiffon. Oxford: Pergamon Press.

Wilkinson, C. R., Nowak, M., Austin, B. & Colwell, R. R. (1981). Specificity of bacterial symbionts in Mediterranean and Great Barrier Reef sponges. *Microb Ecol* **7**, 13-21.

Wilkinson, C. R. (1983a). Net primary productivity in coral reef sponges. *Science* **219**, 410-412.

Wilkinson, C. R. (1983b). Phylogeny of bacterial and cyanobacterial symbionts in marine sponges. *Endocytobiology* **2**, 993-1002.

Wilkinson, C. R., Garrone, R. & Vacelet, J. (1984). Marine sponges discriminate between food bacteria and bacterial symbionts: electron microscope radioautography and *in situ* evidence. *Proc R Soc Lond, B, Biol Sci* **220**, 519-528.

Wilkinson, C. R. (1992). Symbiotic interactions between marine sponges and algae. In *Algae and symbioses: plants, animals, fungi, viruses, interactions explored*, pp. 112-128. Edited by W. Reisser. Bristol: Biopress Ltd.

Williams, S. T., Goodfellow, M., Alderson, G., Wellington, E. M., Sneath, P. H. & Sackin, M. J. (1983). Numerical classification of *Streptomyces* and related genera. *J Gen Microbiol* **129**, 1743-1813.

Williams, S. T., Lanning, S. & Wellington, E. M. H. (1984). Ecology of actinomycetes. In *The biology of the actinomycetes*, pp. 481-528. Edited by M. Goodfellow, M. Mordarski & S. T. Williams. London: Academic Press.

Williams, S. T. (1989). *Bergey's manual of systematic bacteriology*. Baltimore: Williams and Wilkins.

Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V. & Gordon, J. I. (2003). A Genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* **299**, 2074-2076.

Yin, C. & Gao, L. (2000). The microfossils in phosphate depsit in Doushantuo stage, Sinian system, Weng'an, Guizhou province. *Chin Sci Bull* **45**, 279-284.

Yurkova, N., Rathgeber, C., Swiderski, J., Stackebrandt, E., Beatty, J. T., Hall, K. J. & Yurkov, V. (2002). Diversity, distribution and physiology of the aerobic phototrophic bacteria in the mixolimnion of a meromictic lake. *FEMS Microbiol Ecol* **40**, 191-204.

Zaehner, H. & Fiedler, F. P. (1995). Fifty years of antimicrobials: past perspectives and future trends. In *Fifty-third symposium of the Society for General Microbiology*, pp. 67-

84. Edited by P. A. Hunter, G. K. Darby & N. J. Russell. Bath, England, UK.: Cambridge University Press.

Zeng, R., Zhao, J., Zhang, R. & Lin, N. (2005). Bacterial community in sediment from the Western Pacific warm pool. *China Environmental Science* **48**, 282-290.