UNIVERSITY OF CALIFORNIA, SAN DIEGO

Biosynthesis of Scytonemin, A Cyanobacterial Sunscreen

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Oceanography

by

Carla Michelle Sorrels

Committee in charge:

Professor William H. Gerwick, Chair
Professor Kathy Barbeau
Professor Pieter Dorrestein
Professor Tracy Handel
Professor Bradley Moore
Professor Brian Palenik

2009

The Dissertation of Carla Michelle Sorrels is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2009

DEDICATION

My doctoral dissertation is dedicated to my amazing family. Their constant love and support gave me the strength to overcome adversity and never give up. My parents, Joe and Evelyn, are a constant source of love and positive encouragement. They are wonderful people! My future husband, Adam, is truly the light of my life. He makes me smile, keeps me in perspective, and provides unending support no matter what life throws at us. My sisters, Monica and Jeana, are always there to make me laugh and give advice, and my grandparents, Carl and Bette, who provide so much support, kindness, love and generosity. My entire extended family including my grandparents, Jeanine and the late Joe, have always believed in my abilities and given me a source of inner strength. Overall, nothing would be possible without the support of my Lord, Jesus Christ for giving me the hope that nothing is impossible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor William H. Gerwick for his support as the chair of my committee. Through two universities and six years of research, his guidance and love for science have proven to be exceptional qualities in a role model and advisor.

I would also like to acknowledge Dr. Lena Gerwick for her support as a mentor in the laboratory, and the entire Gerwick laboratory for their help and support over the years.

Chapter 2, is being prepared for publication of the material as a review. The dissertation author was the primary investigator and author of this material. I would like to acknowledge Dr. Aishu Ramaswamy for her assistance in learning bioinformatics and RNA isolation techniques in cyanobacteria.

Chapter 3, in part, has been submitted for publication of the material as it will appear in Applied and Environmental Microbiology. 2009. Sorrels, C.M., Proteau, P.J., and Gerwick, W.H.. The dissertation author was the primary investigator and author of this paper.

I would like to acknowledge J. Wells and family for general support to the laboratory.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Sorrels, C.M., Esquenazi, E., Dorrestein, P., and Gerwick, W.H. The dissertation author was the primary investigator and author of this material.

I would like to acknowledge Ed Esquenazi for running the MALDI-TOF instrument used in these analyses and Dr. Dorrestein at University of California, San Diego for the use of the MALDI instrument. I would also like to acknowledge Dr. Kerry McPhail at Oregon State University for assistance in NMR and Dr. Patty Flatt for assistance with initial culturing and extraction of *N. punctiforme* ATCC 29133.

Chapter 5, in part, is currently being prepared for submission for publication of the material. Sorrels, C.M., Suyama, T., and Gerwick, W.H. The dissertation author was the primary investigator and author of this material.

I would like to acknowledge Tak Suyama for providing the synthetic scytonemin monomer substrate and for being a valuable resource of information during the experimentation. I would like to acknowledge Monica Vodrup for assistance with mathematical calculations. I would also like to acknowledge Dr. Alban Pereirabadilla and Dr. Kevin Tidgewell for additional assistance with NMR at Scripps Institution of Oceanography and the University of California, San Diego.

VITA

EDUCATION

Ph.D. Oceanography, Scripps Institution of Oceanography, UCSD. September 2009
Doctoral Candidate, October 2007
Major Field of Study: Biosynthetic Chemistry of Cyanobacteria
Advisor: William H. Gerwick

B.S. Marine Biology, *cum laude*. Texas A&M University at Galveston. May 2003

RESEARCH EXPERIENCE

Postdoctoral Research, University of California, San Diego, San Diego, CA, October 2009
Study of the genetic manipulation of eukaryotic microalgae for the production of therapeutic proteins involved in cancer therapy and for the understanding of primary metabolism important in algal biofuels research
Advisor: Stephen Mayfield

Doctoral Research, Scripps Institution of Oceanography, UCSD, San Diego, CA,
August 2005 – September 2009
Study of the biochemical and genetic elements involved in the biosynthesis of secondary metabolites from cyanobacteria such as scytonemin and hectochlorin
William H. Gerwick. Center for Marine Biomedicine and Biotechnology.

Intern, Cibus LLC, San Diego, CA April 2008 – June 2008
NIH supported internship at the industrial biotechnology company Cibus LLC (San Diego, CA) focusing on plant biotechnology. Research included transposon mutagenesis and the development of mutation selection criteria.

Doctoral Research, Oregon State University, Corvallis, OR, March 2004 – July 2005
Study of the biochemical and genetic requirements for the production of scytonemin, a sunscreen natural product found in cyanobacteria
William H. Gerwick and Philip J. Proteau. College of Pharmacy.

Biological Science Technician. Bureau of Land Management, Corvallis, OR,
June 2003 – August 2003
Employed under the AREMP project to study aquatic and riparian changes in watersheds throughout the Pacific Northwest United States; gained experience in local fish and amphibian identification, laser and computer generated stream surveys, and vegetation and geological surveys.

Research Assistant, Baker Petrolite, Houston, TX, November 2002
Member of a ten person research team studying the use of a chemical to prevent exotic species from spreading between ports through the ballast tanks of ships. Study was based in Venezuela and Florida.

Research Assistant, Dupont, Houston, TX, October 2001 – 2002
Cleaned and sorted samples of marine invertebrates from the Long Island Sound to determine the effectiveness of a remediation project.

Research Assistant, LUMCON, Chauvin, LA, July 2001
Assisted Dr. Nancy Rabalai on a research cruise studying the Dead Zone with the Louisiana Universities Marine Consortium in the Gulf of Mexico; gained hands on experience with research techniques and oceanographic equipment such as flourometers, CTD and Winkler titration.

Research Assistant, Parametrix, Houston, TX, 1999-2000
Employed as a laboratory technician for an ongoing research project on the tributyltin (TBT) content in waters of major shipping channels. Focused on the study of molluscan tissues within the Houston and the Great Lakes shipping channels.

Intern, Pacific Whale Foundation, Maui, Hawaii, March 1998
Internship providing hands-on research experience with the migration habits and potential communication of Humpback whales.

TEACHING EXPERIENCE

Biology Tutor, Tutor.com, Oct. 2006 – Present
Tutor students throughout the United States in all aspects of biology.
Teaching Assistant, Oregon State University, Sept. 2003 – Dec. 2004
Employed by the College of Pharmacy to teach "Terminology of Health Sciences." Solely responsible for lectures, exam evaluations and office hours.

HONORS & AWARDS

National Institutes of Health Pre-doctoral Training Fellowship in Marine Biotechnology. 2007 - 2009
UCSD All-Grad Research Symposium Exemplary Presentation Award. 2008
American Society of Pharmacognosy Graduate Student Travel Grant. 2007
American Society of Pharmacognosy Student Poster Award. 2005
Texas A&M University at Galveston Distinguished Senior Student of the Year. 2003
Texas A&M University at Galveston Aggie Scholar. 2001-2003
Texas A&M University at Galveston Distinguished Sophomore Student of the Year. 2001

PUBLICATIONS

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2009) Organization, evolution and expression of the scytonemin biosynthetic pathway, a cyanobacterial UV-absorbing molecule. *Applied and Environmental Microbiology*. 75: 4861-4869.
Sorrels, C.M., Suyama, T., Proteau, P.J., and Gerwick, W.H. (2009) Probing the enzymatic potential of Scy1263, a hypothetical protein from the scytonemin biosynthetic gene cluster in *Nostoc punctiforme* ATCC 29133. (In preparation).
Sorrels, C.M., Esquenazi, E., Dorrestein, P., Proteau, P.J., and Gerwick, W.H. (2009) Probing the biosynthesis of scytonemin through stable isotope incubation studies (In preparation).
Jones, A.C., Gu, L., Sorrels, C.M., Sherman, D.H., and Gerwick, W.H. (2009) New tricks from ancient algae: Natural product biosynthesis in marine cyanobacteria. *Current Opinion in Chemical Biology*. 13: 216-223.

Gerwick, W.H., Byrum, T., Coates, R.C., Engene, N., Esquenazi, E., Gerwick, L., Grindberg, R., Johnson, M., Jones, A., Malloy, K., Nunnery, J., Pereira, A., Soria, I., Sorrels, C.M., Suyama, T., Tamaguchi, M., Tidgewell, K., Villa, F., Vining, O., Dorrestein, P., Gu, L., and Sherman, D.H. (2008) Integrating chemical and biochemical approaches to natural product drug discovery from marine cyanobacteria. *Conference Proceedings for the International Conference on New Developments in Drug Discovery from Natural Products and Traditional Medicines*.

Gerwick, W.H., Coates, C., Engene, N., Gerwick, L., Grindberg, R., Jones, A.C., and Sorrels, C.M. (2008) Giant marine cyanobacteria produce exciting potential pharmaceuticals. *Microbe*. 3(6): 277-284.

Ramaswamy, A.V., Sorrels, C.M., and Gerwick, W.H. (2007) Cloning and biochemical characterization of the hectochlorin biosynthetic gene cluster from marine cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products*. 70(12): 1977-1986.

Grindberg, R.V., Shuman, C.F., Sorrels, C.M., Wingerd, J., and Gerwick, W.H. (2007) Neurotoxic Alkaloids from Cyanobacteria. *Modern Alkaloids: Structure, Isolation, Synthesis and Biology*. Wiley-VCH Verlag GmbH & Co. KGaA. 139-170.

POSTERS AND PRESENTATIONS

Invited Oral Presentation. American Society of Pharmacognosy Meeting. June 2009.
"Insights into the chemistry and biosynthesis of UV-absorbing natural products from cyanobacteria."

Poster. Society of Industrial Microbiology Annual Meeting. August 2008.
"Biosynthetic Chemistry of Scytonemin: A Cyanobacterial Sunscreen."

Presentation. University of California All-Grad Research Symposium. February 2008.
"Biosynthetic and Chemical Analyses of Cyanobacteria for UV-Absorbing Secondary Metabolites."

Poster. Gordon Research Conference on Marine Natural Products. February 2008.
"Biosynthetic and Chemical Analyses of Cyanobacteria for UV-Absorbing Natural Products."

Poster. American Society of Pharmacognosy Meeting. July 2007.
"Biosynthetic and Chemical Analyses of Cyanobacterial UV-absorbing Secondary Metabolites."

Poster. American Society of Pharmacognosy Meeting. July 2005.
"Biosynthesis of scytonemin: Cyanobacterial Sunscreen."

Poster. Center for Gene Research and Biotechnology Retreat. September 2004.
"Scytonemin - Nature's Cyanobacterial Sunscreen, Search for the Genetic and Biosynthetic Pathway."

Poster. Volcano Conference. February 2005.
"Scytonemin – Nature's Cyanobacterial Sunscreen, Pursuit of the Genetic and Biosynthetic Pathway."

OTHER RESEARCH RELATED EXPERIENCE

Advanced/Rescue Diver, PADI certified, AAUS Scientific Certification
First Aid and CPR Certified
Texas Boaters Education License

ACADEMIC SERVICE

Sea Grant Expanding Students Horizons Volunteer and Lab Coordinator. 2008-2009
Texas A&M University at Galveston Campus Judicial Board Appointment. 2001-2003
Texas Academy of Science, Galveston Branch, President. 1999-2002

PUBLIC SERVICE

Women in Bioscience and Technology Conference Planning Committee. San Diego, CA. 2008-2009.
St. Mark's United Methodist Church Young Adult Group Leader. San Diego, CA. 2007-Present
St. Mark's United Methodist Church Youth Boardmember. San Diego, CA. 2006-Present
YMCA Executive Board Secretary. Galveston, TX. 2002
YMCA Board of Directors. Galveston, TX. 2001-2003

ABSTRACT OF THE DISSERTATION

BIOSYNTHESIS OF SCYTONEMIN, A CYANOBACTERIAL SUNSCREEN

by

Carla Michelle Sorrels

Doctor of Philosophy in Oceanography

University of California, San Diego, 2009

William H. Gerwick, Chair

Depletion of the stratospheric ozone layer is increasing the level of harmful ultraviolet (UV) radiation reaching the earth's surface. The environmental impacts of this radiation are largely unknown; however UV-induced DNA damage may lead to effects on primary productivity in plants and the prevalence of skin cancer. Cyanobacteria are photosynthetic prokaryotes with an evolutionary history that precedes the development of atmospheric ozone protection. These factors make cyanobacteria a valuable resource for the study of adaptations to UV radiation,

including the production of UV absorbing secondary metabolites. One of these metabolites is scytonemin, a yellow-brown pigment found in the sheaths of many cyanobacteria. Its unique dimeric indolic-phenolic structure, powerful UV-A absorbing properties, and biological activity in biomedical assays prompted our study of its biosynthesis.

This thesis involved a multidisciplinary investigation of the biosynthesis of scytonemin. Preliminary evidence for genes involved in scytonemin biosynthesis were identified through bioinformatic analyses and semi-quantitative reverse transcriptase PCR using the *N. punctiforme* ATCC 29133 genome. These genes are shown to be a part of a transcriptional gene cluster that is upregulated after exposure to UV radiation. Examination of this gene cluster across cyanobacterial lineages reveals unique genetic characteristics in individual clusters, and further suggests ancient evolutionary history for its biosynthesis.

Matrix Assisted Laser Desorption-Time of Flight (MALDI-TOF) mass spectrometry technique was utilized in a novel method to understand the biosynthetic precursors used for scytonemin biosynthesis through stable isotope incubation studies. The MALDI-TOF technique also allowed a unique glimpse into the near-real-time induction and rate of scytonemin biosynthesis. The final study reported in this thesis reveals the function of a gene uniquely present in the scytonemin gene cluster of *N. punctiforme* ATCC 29133. This gene is shown to encode a protein with *in vitro* oxygenase activity, similar to the well characterized mushroom tyrosinase.

BIOSYNTHESIS OF SCYTONEMIN, A CYANOBACTERIAL SUNSCREEN

CHAPTER ONE -

GENERAL INTRODUCTION

HIGHLIGHTS OF NATURAL PRODUCTS AND THE EVOLUTION OF CYANOBACTERIA IN RELATION TO ULTRAVIOLET RADIATION

History of Natural Products

The use of natural substances for medicinal purposes is described as early as 78 A.D. when Dioscorides wrote of the thousands of plants used for medicinal purposes in "De Materia Medica" (Clark, 1996). Today, these natural substances called "natural products" have grown to be a significant source of new drugs used against many types of diseases (Butler, 2005). The discovery, development, and classification of drugs such as the anti-malarial quinine **(1)**, the antibiotic penicillin G **(2)**, and the analgesic morphine **(3)** revolutionized the use of natural products against specific human diseases and made them a cornerstone of modern medicine (Clark, 1996). Examples of historically important natural products are shown in **figure 1.1**. Today, over 70% of medicinal drugs are inspired by natural sources (Newman and Cragg, 2007).

Despite the use of natural products throughout history, the biological basis for the effect of these metabolites on human disease did not undergo scientific scrutiny until the mid-1800's when the pharmacologically active cardiac glycosides were discovered. One of these glycosides, digoxin **(4)**, was isolated from the foxglove plant *Digitalis purpurea* and found to exert powerful and selective action on cardiac muscle, and became important in the treatment of heart disease. The usefulness of natural products was further strengthened by the early discovery of bioactive alkaloids such as morphine and atropine **(5)** (Clark, 1996). Morphine is a metabolite of the opium poppy that acts as a potent analgesic used today to treat moderate to severe pain, while atropine is derived from the belladonna plant and is used as a smooth muscle relaxant (Brielmann *et al.*, 2006). Finally, in 1928 observations by Alexander Fleming changed the face of modern medicine and the field of natural products with the discovery of the penicillins, secondary metabolites produced by the fungus *Penicillium*. This discovery started the golden era of antibotics in medicine and increased the awareness and desire for natural products research (Fleming, 1944; Clark, 1996).

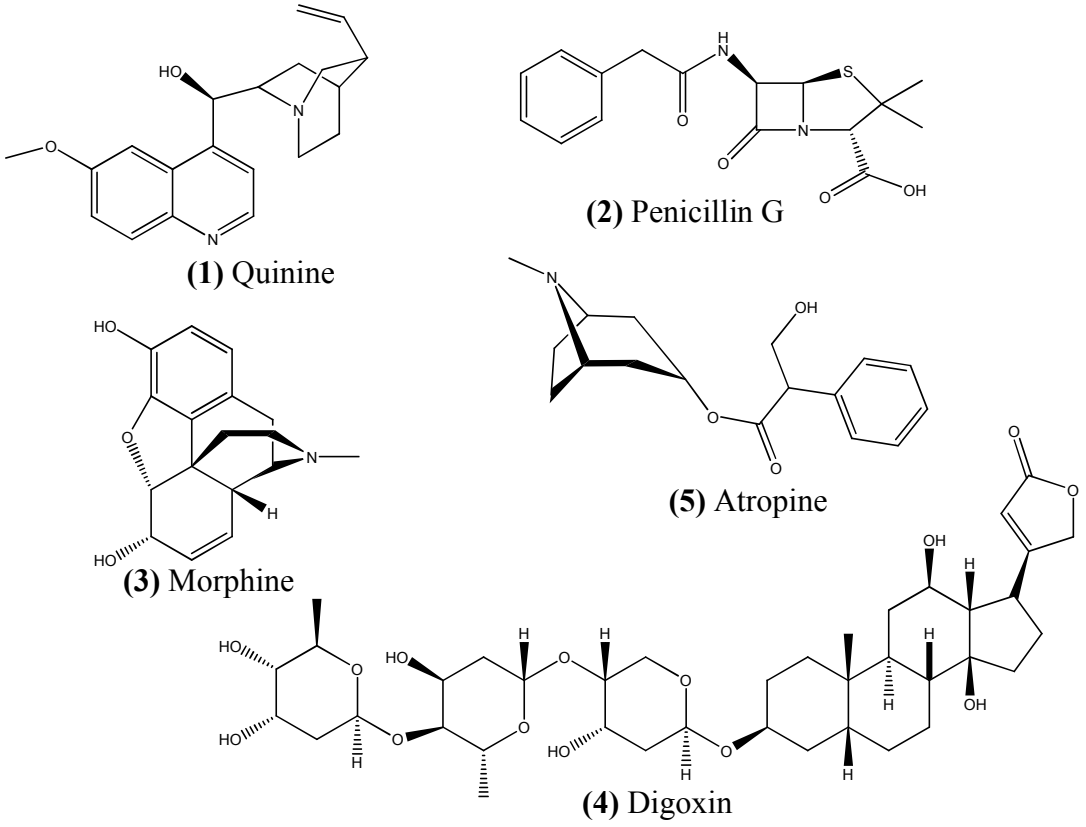


Figure 1.1: Chemical structures of historically and biomedically relevant natural products

In the 1970's, scientists began to exploit a new frontier for natural products research: the ocean. The oceans cover about 70% of the Earth and are estimated to account for between 1 and 2 million species of marine organisms (Simmons *et al*., 2005). Much of this rich species diversity is compacted into coastal environments. The impact of a large number of marine organisms living in a small region leads to a complex and competitive ecosystem. Thus, many marine organisms from these

regions produce toxic secondary metabolites that function in defense against predation, disease, or overgrowth (Albrizio *et al*., 1995; Kubanek *et al*., 2002; Simmons *et al.*, 2005). The complexity of these interactions and the types of metabolites produced by these organisms led to the development of three parallel tracks in marine natural product research: marine toxins, marine biomedicines, and marine chemical ecology (Faulkner, 2000). Marine natural products research has led to the discovery of a wide range of molecules with unprecedented molecular structures and potent bioactivities. These molecules have provided valuable insight into the ecological, biotechnological and pharmacological importance of unique organisms such as the cyanobacteria.

Cyanobacterial Natural Products

Cyanobacteria are gram-negative prokaryotes thought to be among the most ancient organisms on the planet. They use photosystems I and II to carry out oxygenic photosynthesis similar to plants (Cohen and Gurevitz, 2006). Their ability to use light to create energy through photosynthesis and their wide global distribution in diverse aquatic, marine and terrestrial environments have led to cyanobacteria becoming one of the most important primary producers on the planet (Tomitani *et al.*, 2006). Cyanobacteria also play a number of other ecological roles that greatly impact their surrounding environments and interactions with other organisms, including providing a source of nitrogen through symbiotic interactions with higher plants, the formation of dense toxic blooms known as Harmful Algal Blooms (HABs) and the formation of many biologically active secondary metabolites (Whitton and Potts, 2000).

Cyanobacterial secondary metabolites can be broadly classified into two categories in natural product research: (i) biotoxins and (ii) biomedicines including cytotoxins and other bioproducts (Burja *et al.*, 2001; Jaiswal *et al.*, 2008). Biotoxins cause acute and often lethal poisoning and have been implicated in human and livestock deaths, which normally result from cyanobacterial ingestion during bloom events (Burja *et al.*, 2001; Carmichael *et al*., 2001). These toxic cyanobacteria blooms were first reported in 1878 when *Nodularia spumigena* found in Australia was reported to poison both sheep and cattle (Carmichael, 2008). Due to their risks for serious health and economic problems, cyanobacterial blooms and biotoxins have been intensely studied. Freshwater blooms typically produce two types of biotoxins: (i) hepatotoxins and (ii) neurotoxins. Hepatotoxins are the most commonly encountered toxins in cyanobacteria and include the microcystins **(6)** from many cyanobacterial genera including *Microcystis*, nodularins **(7)** from *Nodularia*, and cylindrospermopsin **(8)** from *Cylindrospermopsis*. Cyanobacterial neurotoxins include anatoxin-a **(9)** from *Anabaena* and *Oscillatoria*, anatoxin-a(s) **(10)** from *Anabaena flos-aquae*, and saxitoxin **(11)** from *Aphanizomenon* and *Anabaena* (Namikoshi and Rinehart, 1996; Grindberg *et al*., 2007). Biotoxins are produced by cyanobacteria in both terrestrial and marine environments. β-N-methyl-amino-L-alanine (BMAA, **12**) is an example of a terrestrial biotoxin. This toxin is produced by the genus *Nostoc*, a genus known to have a symbiotic relationship with the cycads trees present in Guam. Through biomagnification, BMAA is thought to play a significant role in the high incidence of the deadly neurodegenerative disease amyotrophic lateral sclerosis/parkinsonism

dementia complex (ALS/PDC) in humans (Grindberg *et al*., 2007; Jonasson *et al*., 2008). Another example of a cyanobacterial biotoxin that is found in the marine environment is lyngbyatoxin A from the Hawaiian strain of *Lyngbya majuscula* **(13)**. Lyngbyatoxin A is the causative agent for a common blistering dermatitis called "swimmer's itch" in Hawaiian waters and has been implicated in several poisonings due to the consumption of turtle meat (Tidgewell *et al.*, 2009). Chemical structures for these cyanobacterial biotoxins are shown in **figure 1.2**.

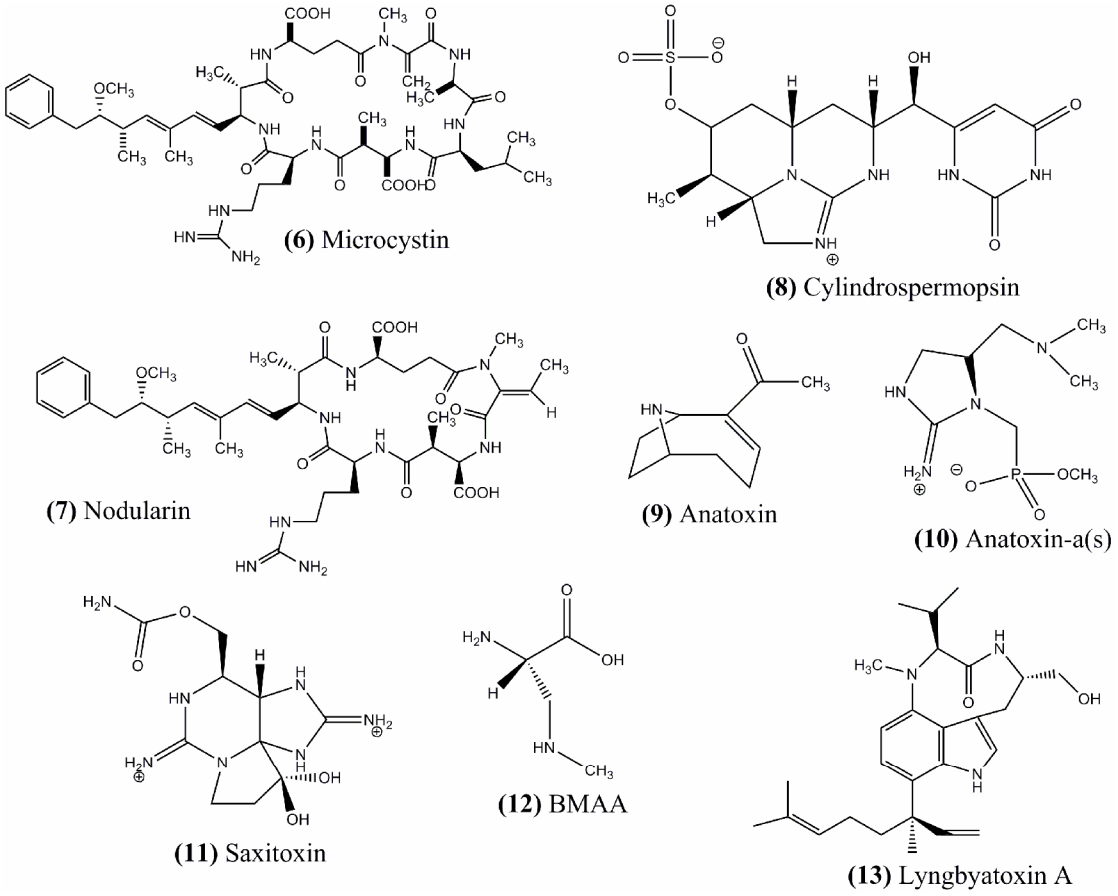


Figure 1.2: Chemical structures of cyanobacterial biotoxins

Despite their reputation for the production of environmentally toxic compounds, cyanobacteria (particularly those from the marine environment), also produce a broad range of metabolites with medically relevant bioactivities. These metabolites have been found to function as anticancer, antibiotic, antiviral, antifungal, and anti-inflammatory agents. The medicinal potential of cyanobacteria was first realized in the 1500s when they were used to treat gout, fistula, and cancer (Burja *et al.*, 2001). However, the true potential of marine cyanobacteria as prolific producers of bioactive natural products was not understood until the efforts of Richard E. Moore (University of Hawaii) brought these organisms into the spotlight (Tidgewell *et al.*, 2009). A majority of the secondary metabolites isolated from marine cyanobacteria contain nitrogen and come from the order Nostocales, particularly the genera *Lyngbya*, *Oscillatoria*, and *Symploca* (Burja *et al.*, 2001; Tan, 2007). The need for cyanobacteria to defend themselves in a highly competitive marine environment has resulted in the diversity of these compounds that affect numerous targets in eukaryotic cells (Burja *et al.*, 2001). Many cyanobacterial metabolites target tubulin or actin in eukaryotic cells, leading to an interest in these molecules as anti-cancer agents (Tan, 2007). Two of these anti-tubulin metabolites from the marine environment are curacin A **(14)** isolated from a Caribbean collection of *L. majuscula* and dolastatin 10 **(15)** originally isolated from the sea hare *Dolabella auricularia* and later found to be produced by the cyanobacterium *Symploca* sp. VP642 from Palau (Pettit *et al.*, 1987; Bai *et al.*, 1990; Gerwick *et al.*, 1994; Luesch *et al.*, 2001). These molecules have also served as pharmacological leads for the development of synthetic analogs used in

clinical trials such as the cryptophycin analog TZT-1027 and curacin A analog 3,4,5-trimethoxybenzaldehyde O-((2E,4E,8RS)-8-hydroxy-5-methyl-8-thiophen-2-ylocta-2,4-dienyl)oxime (Wipf *et al.*, 2002; Tan, 2007). Another cyanobacterial molecule with potential as an anticancer agent is cryptophycin 1 **(16)**. The cryptophycins were isolated from *Nostoc* sp. ATCC 53789 and *Nostoc* sp. GSV224 and are known to target microtubules, particularly the Vinca site of tubulin (Kerksiek *et al.*, 1995). This natural product was also found to be particularly effective against multiply-drug resistant ovarian carcinoma and breast carcinoma cells (Smith *et al.*, 1994). Analogs of cryptophycin are actively being explored in clinical trials towards resistant forms of ovarian cancer (D'Agostino *et al.*, 2006).

Many cyanobacterial metabolites have also been found to have potent neurotoxic activity (Tan, 2007). Kalkitoxin **(17)** and jamaicamide A **(18)** from the marine cyanobacterium *L. majuscula* have exhibited activity as voltage-gated sodium channel blockers (Edwards *et al.*, 2004; LePage *et al.*, 2005). Kalkitoxin has been shown to block veratridine induced neurotoxicity and to be a potent inhibitor of the elevation of intracellular Ca^{2+} concentration that can accompany neurons after exposure to veratridine (LePage *et al.*, 2005). Another neurotoxin from *L. majuscula*, antillatoxin **(19)**, works in an opposite manner by increasing the intracellular Ca^{2+} concentration; therefore, acting as an activator of the eukaryotic voltage-gated sodium channels (Li *et al.*, 2001). Cyanobacterial metabolites are not only active as neurotoxins and anti-proliferative agents but have a variety of other diverse bioactivities. These include activities as virucidal agents such as cyanovirin from

Nostoc ellipsosporum, antifungal agents such as hectochlorin from *L. majuscula*, and anti-inflammatory agents such as malyngamide S from *Lyngbya* sp. (Boyd *et al.*, 1997; Marquez *et al.*, 2002; Tan, 2007).

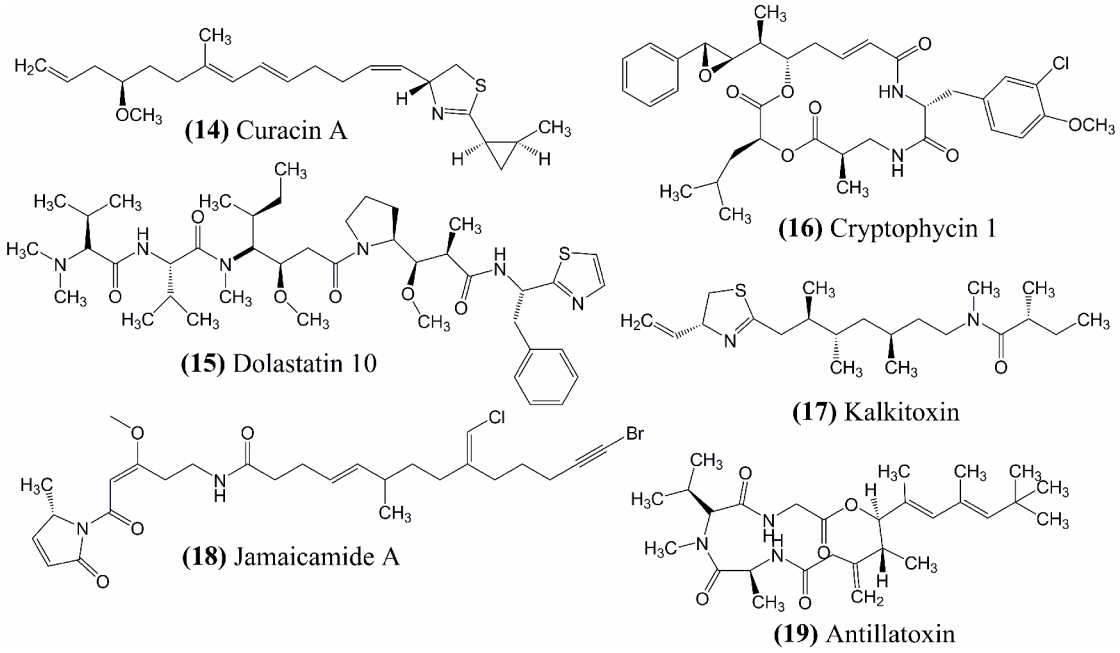


Figure 1.3: Chemical structures of cyanobacterial cytotoxic and bioactive metabolites

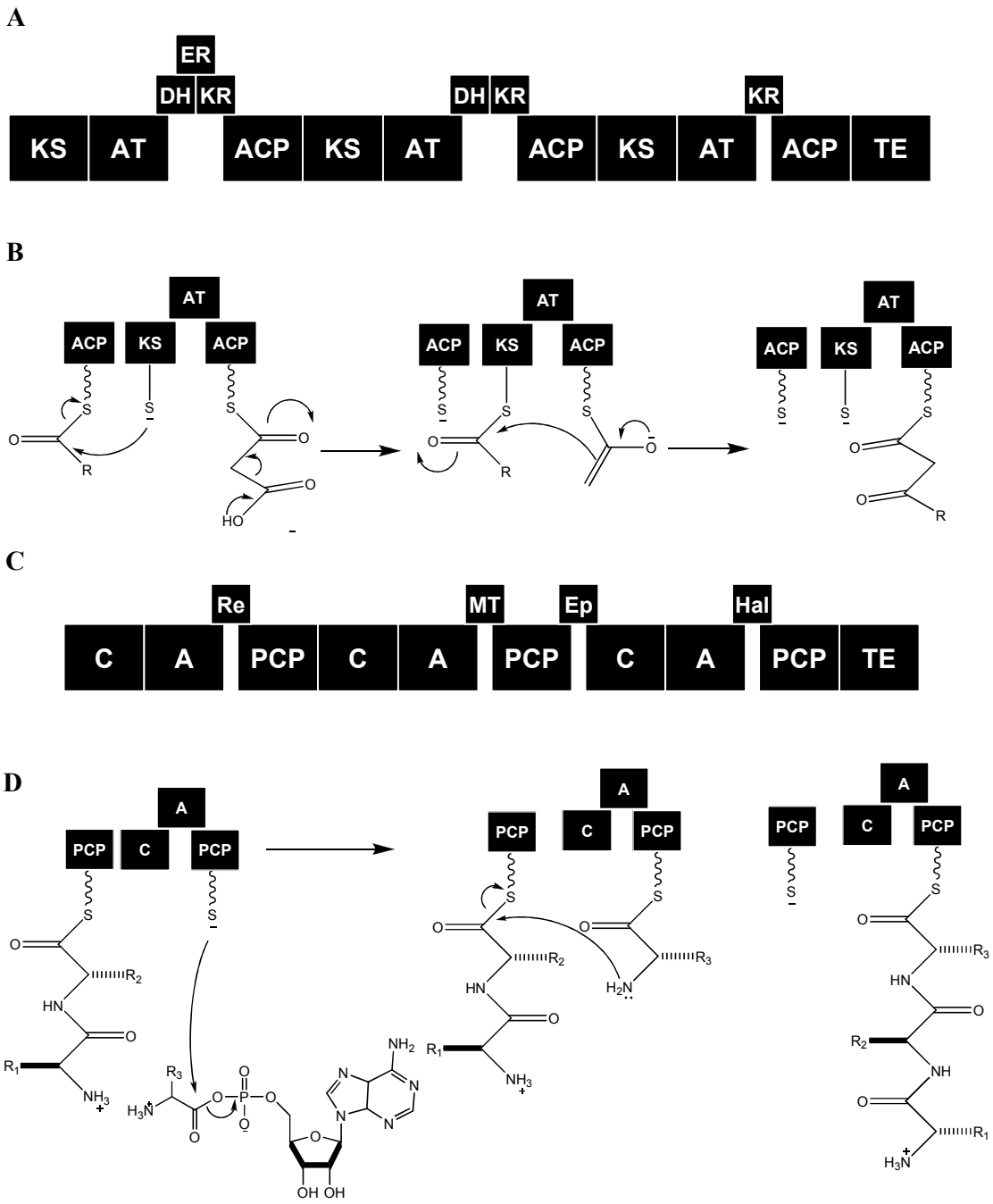
Cyanobacterial Natural Product Biosynthesis

The presence of the prolific combinations of structural units in cyanobacterial toxins and bioactive natural products is a witness to their large biosynthetic potential. Cyanobacteria harbor a wide array of unique and novel enzymatic mechanisms leading to the high structural diversity of their secondary metabolites (Dittmann *et al.*, 2001). Many of the bioactive secondary metabolites produced by cyanobacteria are biosynthesized through the step-by-step incorporation of individual building blocks in pathways containing multimodular proteins known as polyketide synthases (PKS) and

nonribosomal peptide synthetases (NRPS). Polyketide synthases typically incorporate monoacyl thioesters using a series of core reactions catalyzed by an acyltransferase, a ketosynthase and an acyl-carrier protein. These acyl groups can be variably reduced using accessory domains consisting of ketoreductases, dehydratases and enoylreductases. Nonribosomal peptide synthetases are similar in genetic architecture and chemical mechanism to polyketide synthases in that they incorporate amino acid type molecules using three core domains: an adenylation domain, a condensation domain, and a peptidyl-carrier protein. Both types of biosynthetic pathways are typically released via a thioesterase domain that can result in either linearized or cyclic natural products (Fischbach and Walsh, 2006). The general architecture of these pathways and the basic chemical reactions catalyzed by the core domains are diagrammed in **figure 4**.

Cyanobacteria have used the versatility of these biosynthetic pathways to greatly diversify their secondary metabolites providing a variety of interesting chemical structures of biotechnological and pharmacological potential. However, the biosynthetic diversity of cyanobacteria is not limited to NRPS and PKS pathways. Cyanobacteria also use pathways similar to those in primary metabolism to produce secondary metabolites. These metabolites can be used as adaptations to certain spectral ranges of intense light such as the use of scytonemin to screen ultraviolet (UV) radiation.

Figure 1.4: Diagram of the architecture and condensation mechanism for PKS and NRPS biosynthetic pathways. A) Basic domain architecture of the PKS pathway including some common accessory domains, B) Condensation mechanism of PKS, C) Basic domain architecture of NRPS pathway including some common accessory domain, and D) Condensation mechanism of NRPS.



Domains

KS - Ketosynthase
AT - Acyltransferase
ACP - Acyl-carrier-protein
DH - Dehydratase
KR - Ketoreductase
ER - Enoylreductase

C - Condensation
A - Adenylation
PCP - Peptidyl-carrier-protein
Re - Reductase
MT - Methyltransferase
Ep - Epimerase
Hal - Halogenase

Evolution and Cyanobacteria

The production of secondary metabolites in cyanobacteria as adaptations to high light intensity, particularly wavelengths in the UV region of the electromagnetic spectrum, played a crucial role in their evolution (Wynn-Williams *et al.*, 2002). The presence of filamentous stromatolitic microfossils found in the Apex chert deposits in Western Australia indicate that ancient cyanobacteria known as protocyanobacteria inhabited the earth over three billion years ago (Garcia-Pichel, 1998; Paul, 2008). These stromatolitic formations were probably analagous to present day cyanobacterial mats found in hypersaline lagoons, hot springs and other restricted environments (Garcia-Pichel, 1998).

Protocyanobacteria are thought to have evolved to form ancestral cyanobacteria resembling the modern phylum over the next 500 million years. An increased environmental presence of cyanobacteria around 2.5 billion years ago is supported by geological surveys of the sedimentary biomarker 2α-methylhopane. This biomarker is related to 2-methylbacteriohopanepolyols, a metabolite produced by present day cyanobacteria and few other organisms (Summons *et al.*, 1999; Des Maris, 2000; Paul, 2008). The structural similarities between these molecules are apparent in **figure 1.5**.

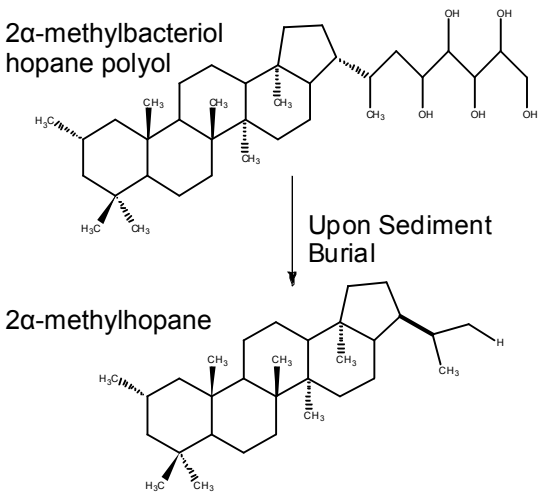


Figure 1.5: Chemical structures for 2α-bacteriolhopane polyl and the 2α-methylhopane biomarker.

During this period of evolution, cyanobacteria developed the capability of oxygenic photosynthesis. Oxygenic photosynthesis uses a specialized photosynthetic apparatus consisting of two photosystems linked in a series, photosystem I and photosystem II. Photosystem II carries out water oxidation by splitting water into molecular oxygen, protons, and electrons driven by light energy. Photosystem I produces ATP through the reduction of ferredoxin to drive metabolism (Blankenship and Hartman, 1998; Dismukes *et al.*, 2001). The geologic record suggests that oxygenic photosynthesis first evolved prior to 2.8 billion years ago at relatively the same time as the appearance of cyanobacteria (Des Marais, 2000; Paul, 2008). The photosystems developed during this period of geologic history were later incorporated into eukaryotic cells as the chloroplast, giving rise to eukaryotic algae and higher plants (Dismukes *et al.*, 2001). The formation of molecular oxygen through oxygenic

photosynthesis probably increased global productivity by at least two to three orders of magnitude and played a critical role in oxygenation of the atmosphere allowing for the development of higher life forms (Des Marais, 2000; Whitton, 2000; Dismukes *et al.*, 2001). A timeline of cyanobacterial evolution in relation to changes in atmospheric oxygen is diagrammed in **figure 1.6**.

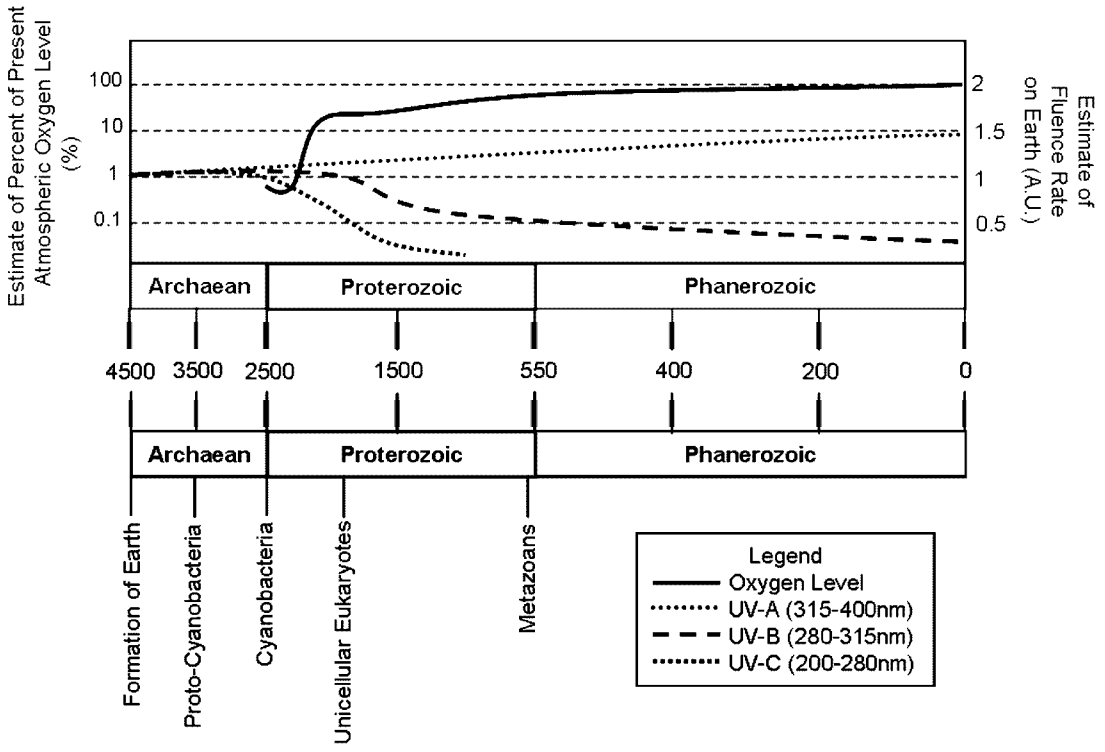


Figure 1.6: Comparison of atmospheric oxygen levels with ultraviolet radiation fluence rates on the Earth over geologic history labeled with documented fossil records. Figure derived from Schopf, 1994; Garcia-Pichel, 1998; Berman-Frank *et al.*, 2003.

The development of oxygen on Earth not only played a major role in the formation of higher life forms, but also on atmospheric geochemistry. Prior to atmospheric oxygen, the Earth was bombarded by unfiltered wavelengths of UV radiation and visible light. Although visible light was important for primary production and the biogenesis of oxygen, UV radiation was dangerous for living organisms. UV radiation is split into three ranges based on the energy level of its photons. The first range, UV-C (200-280nm) has the lowest wavelengths in the UV region that correspond to the highest photon energy levels and is extremely harmful due to its interactions with proteins and DNA. UV-B (280-315nm) also targets DNA and can inhibit various biological activities such as nitrogen fixation, ATP synthesis, and the synthesis of chlorophyll a. The region of lowest energy, UV-A (315-400nm), is the most prevalent form of UV radiation reaching the Earth today, and is associated with the production of reactive oxygen species and the inhibition of photosynthesis (Castenholz and Garcia-Pichel, 2000).

As oxygen built up in the geologic atmosphere, it began to react with lightening and high energy UV-C radiation to form a new molecule, ozone (Garcia-Pichel, 1998; Rowland, 2006). The build up of ozone in the atmosphere formed the ozone layer. This atmospheric layer is located 25-30 kilometers above the surface of the planet in the stratosphere (Parson, 2003). Both oxygen and ozone can absorb these short wavelength high energy photons preventing solar radiation less than 290 nm from penetrating the stratosphere. The ozone layer acts as a blanket around the Earth,

preventing biological damage due to UV radiation exposure, and plays a role in stabilizing the balance of heat in the atmosphere (Rowland, 2006). The evolution of cyanobacteria prior to the formation of the protective stratospheric ozone layer required that they handle a difficult ecological challenge due to the harsh UV-A, UV-B, and UV-C radiation. In order to combat this challenge, cyanobacteria developed multiple UV defense mechanisms. Although the present day ozone layer screens most of the UV radiation, these mechanisms continue to protect cyanobacteria and many higher organisms from the lower amounts of UV radiation that are present in the modern environment (Ehling-Shultz and Scherer, 1999).

Ultraviolet Radiation Defense Mechanisms in Cyanobacteria

One of the common UV radiation defense mechanisms in cyanobacteria is avoidance. Although some modern cyanobacteria permanently live in shaded environments, many are known to be capable of moving or gliding in their natural environments (Ehling-Schulz and Scherer, 1999). These motions often correspond with a reaction to a stimulus, particularly light levels, causing cyanobacteria to move vertically in sediments, water columns, and cyanobacterial mats to reach optimal photosynthetic intensities (Stal, 2000). The mechanism of avoidance probably evolved to combat the high levels of UV radiation during early anoxic atmospheric conditions. The presence of reduced iron [Fe(II)] in solution in the ocean, as supported by the presence of banded iron formation in the stratigraphic record, would have provided a significant UV-C screen. The ability to move in the water column

allowed cyanobacteria to harness photosynthetic light while using inorganic elements to screen the UV radiation (Olson and Pierson, 1986). Avoidance is also a mechanism commonly used in microbial mat ecosystems. Microbial mats are multilayered microbial communities growing in a variety of environments. Cyanobacteria are particularly efficient at mat formation due to their photosynthetic and nitrogen fixation capabilities. Generally, microbial mats form vertically stratified layers containing different microorganisms within each layer based on their nutrient and light requirements. This layering often consists of sand or scytonemin on top followed by cyanobacteria. Research has shown that cyanobacteria present in this second layer migrate up and down in the mat in a daily manner and undergo significant photoinhibition in the absence of migration (Stal, 2000). The ability to descend in the mat by 100-500 micrometers allows cyanobacteria to increase their photosynthetic efficiency (Kruschel and Castenholz, 1998; Stal, 2000).

The end of banded iron formation indicates that a process of iron oxidation resulted in the precipitation of insoluble ferric minerals. This precipitation limited the supply of reduced iron in the waters and resulted in an increased transparency of these waters to UV radiation (Garcia-Pichel, 1998). The need for further defenses against this radiation may have played a role in the ability of cyanobacteria to actively correct lesions caused by UV radiation through the repair of DNA. The early evolution of DNA repair in response to UV radiation is supported by these mechanisms predating the diversification of higher organisms due to the presence of homologous mechanisms in prokaryotes, archaea, and eukaryotes (Woese, 1987). The two main

types of lesions caused by UV radiation are cyclobutane pyrimidine dimers (CPD) and (6-4) pyrimidine-pyrimidone photoproducts [(6-4)PP] diagrammed in **figure 1.7**.

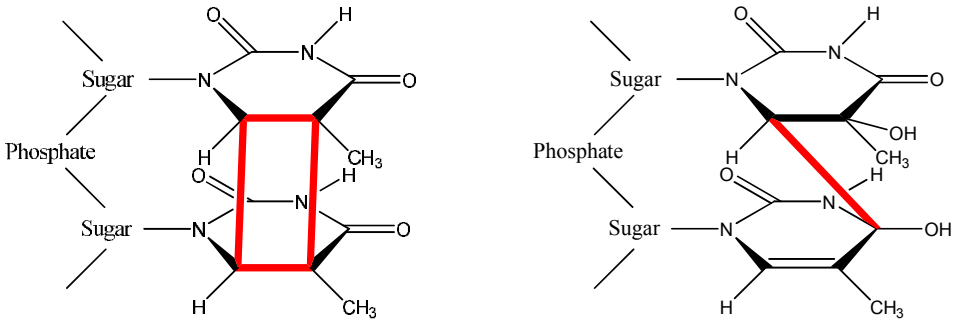


Figure 1.7: Diagram of the chemical structure of DNA lesions caused by UV radiation

These lesions usually occur through the dimerization of adjacent pyrimidines and directly affect DNA function, often interfering with primary biological processes leading to a loss of function in the cell. Cyanobacteria are known to be capable of correcting these lesions through two major mechanisms, photoreactivation and nucleotide excision repair. Photoreactivation occurs when a photolyase enzyme containing the cofactors diazaflavin and flavin adenine dinucleotide (FAD) uses the energy provided by blue light to cleave the DNA lesion (Castenholz, 2000). Nucleotide excision repair occurs when the UvrABC endonuclease is activated through the SOS response and binds to a DNA lesion. A strand of 12 nucleotides is

then excised from the DNA and removed. Once removed, DNA polymerase I uses the complementary strand to resynthesize the missing DNA strand without the original lesion (Snyder and Champness, 2003).

Despite the presence of DNA repair mechanisms, the effects of UV radiation on cyanobacteria would still be detrimental without an additional method of protection. Both the photosynthetic apparatus and nitrogenase, two of the most important evolutionary advances from cyanobacteria, have been found to be extremely susceptible to damage by UV radiation (Sinha *et al*., 1996). The evolution of secondary metabolites capable of absorbing UV radiation provided cyanobacteria with a means of UV protection while maintaining exposure to visible light.

Ultraviolet Radiation Stimulated Secondary Metabolites in Cyanobacteria

The photosynthetic nature of cyanobacteria explains their production of many pigments including chlorophyll a, phycoerythrin and phycocyanin. These pigments are designed to capture photosynthetically active light and hence have a maximum absorption ranging from 400-700 nm (Mathews *et al.*, 2000). However, these photosynthetically active pigments do not provide protection from harmful UV radiation; therefore, cyanobacteria also produce other secondary metabolites to capture radical oxygen species and to absorb light in the lower wavelengths. **Figure 1.8** diagrams the light absorbance profile of various pigments commonly found in cyanobacteria.

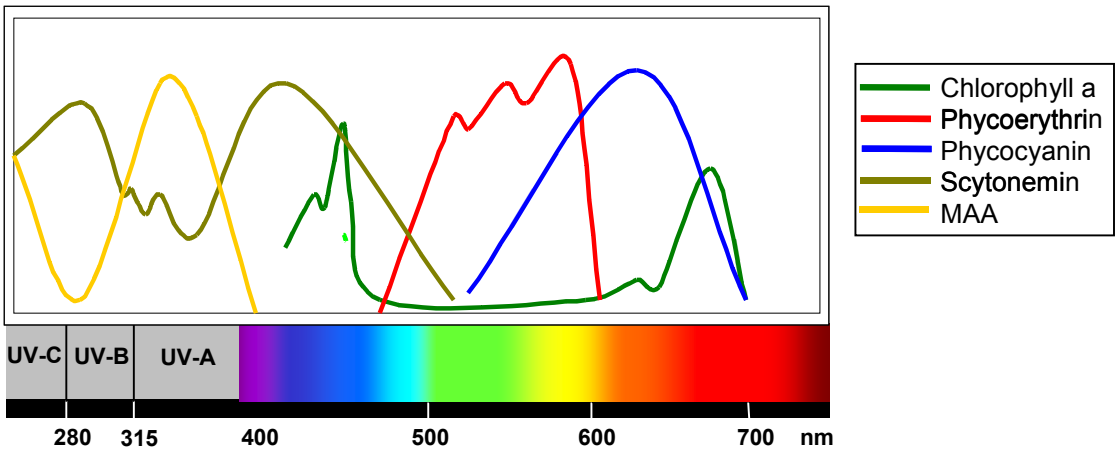


Figure 1.8: Diagrammed absorbance profile of common visible light and UV radiation absorbing molecules found in cyanobacteria. Spectra adapted from Cockell and Knowland, 1999; Proteau, *et al*., 1999; Mathews, *et al*., 2000.

UV absorbing compounds in cyanobacteria usually consist of molecules with extensive π-electron systems. These systems absorb excess energy through electronic transitions known as $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions. These transitions occur when electrons involved in multiple bonding (π) or nonbonding heteroatom electrons (n) absorb the excess light energy causing a transition from lower to higher energy states (Crews *et al.*, 1998). The absorbance properties of UV absorbing molecules can be altered by the size of the molecule, aromaticity, substitution pattern, and the increasing complexity of a π-electron system. Examples of UV absorbing molecules include the mycosporine amino acids found in various organisms, scytonemin found in cyanobacteria, the phenylpropanoids of plants, and the melanins found in bacteria, fungi and animals (Cockell and Knowland, 1999).

Mycosporine-like amino acids (MAAs) are water soluble, shikimate pathway derived substituted cyclohexanone or cycloheximine molecules linked to amino acids or imino alcohols that absorb UV radiation (Bandaranayake, 1998). The structurally simplest MAAs are found in various forms of life including cyanobacteria, eukaryotic micro-algae and fungi, which supports the presence of these molecules early in evolution. The modern suite of MAAs have a maximum absorbance in the UV-B range between 310 and 360 nm depending on the attached amino acid residue (Garcia-Pichel, 1998). Questions still remain as to the true nature of the UV screening capacity of MAAs in the modern environment. However, in the cyanobacterium *Nostoc commune* two out of three photons are absorbed before reaching the intracellular space when MAAs are present in the extracellular glycan (Ehling-Schulz and Scherer, 1999). There are 19 identified MAAs found in numerous marine organisms ranging in taxonomy from cyanobacteria to species of Antarctic fish. Some of the common MAAs are diagrammed in **figure 1.9**. Many higher organisms have multiple types of MAAs that are believed to accumulate from the diet. The presence of more than one type of MAA in a single organism is thought to provide a broader range of UV protection due to the slightly different absorbance maximum for each MAA (Cockell and Knowland, 1999).

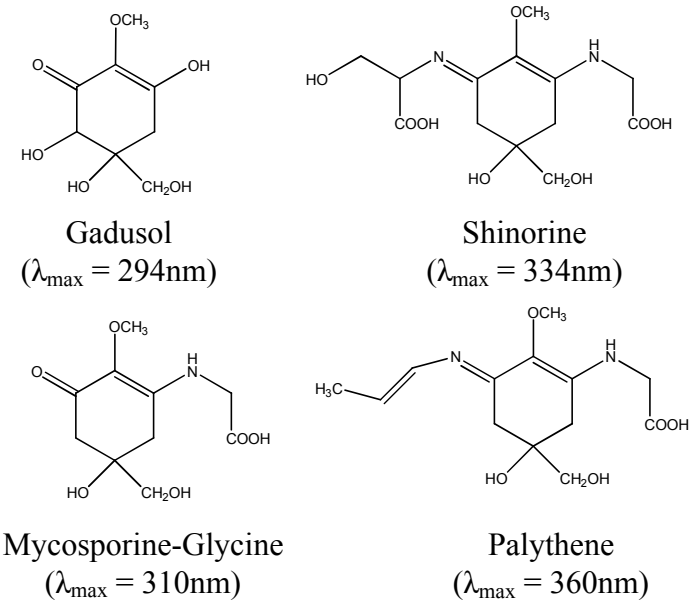
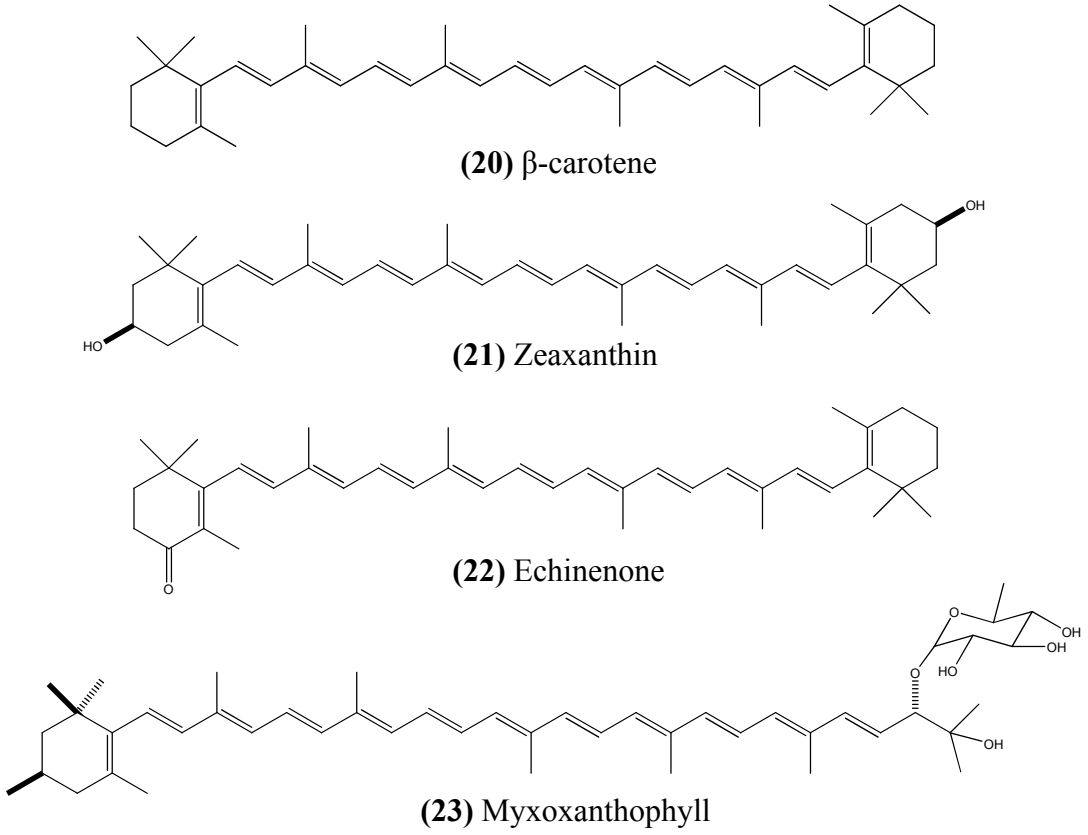


Figure 1.9: Chemical structures of mycosporine amino acids found in cyanobacteria.

As photosynthesis increased the levels of oxygen in the atmosphere, cyanobacteria would have been one of the first organisms to deal with the harmful effects of the third level of UV radiation, UV-A (Garcia-Pichel, 1998). UV-A is associated with the production of reactive oxygen species when molecular oxygen is present (Castenholz and Garcia-Pichel, 2000). One type of molecule produced by many organisms in response to these reactive oxygen molecules are the carotenoids. As the largest class of naturally occurring pigments, carotenoids are responsible for most of the many shades of yellow, orange and red found in microorganisms, fungi, algae, plants, and animals. These pigments typically consist of a C_{40} hydrocarbon backbone with a series of conjugated double bonds that are produced from eight C_5

isoprenoid units derived from the general isoprenoid biosynthetic pathway (Hirschberg and Chamovitz, 1994). The carotenoids commonly found in cyanobacteria include β-carotene **(20)**, its hydroxyl derivatives such as zeaxanthin **(21)**, its keto derivatives such as echinenone **(22)**, and carotenoid glycosides such as myxoxanthophyll (**23**). The composition of these carotenoids is dependent on growth conditions such as growth stage, light intensity, nitrogen source and concentration as well as the individual strains being tested (Takaichi and Mochimaru, 2007).



(20) β-carotene

(21) Zeaxanthin

(22) Echinenone

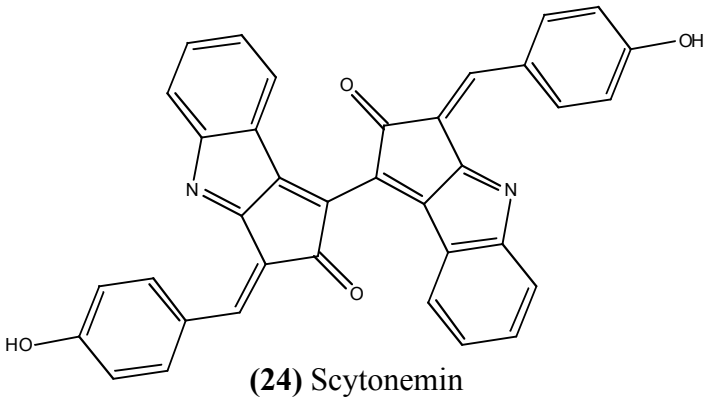
(23) Myxoxanthophyll

Figure 1.10: Chemical structures of carotenoids commonly found in cyanobacteria.

In all photosynthetic organisms, carotenoids are known to serve two major functions: accessory pigments for light harvesting and prevention of photooxidative damage. Their role in photooxidative damage stems from the ability of carotenoids to remove singlet oxygen, triplet chlorophyll, and inhibit lipid peroxidation (Ehling-Schulz and Scherer, 1999). Cyanobacteria use carotenoids to protect themselves from increased oxidative stress incurred during exposure to sunlight for photosynthesis. Their exposure to UV-A radiation from the sun can generate reactive oxygen intermediates. These toxic reactive oxygen intermediates can lead to DNA mutation (Ehling-Schulz and Scherer, 1999). Although carotenoids are extremely important as a first line response to UV radiation acclimation, they have never demonstrated significant UV screening advantages due to an absorbance typically far away from the UV ranges (e.g. in the visible light range; Cockell and Knowland, 1999).

As the evolution of cyanobacteria continued and they began to spread to terrestrial environments, cyanobacteria needed an even greater defense against UV radiation. Optimally, this defense would not require a constant influx of energy. They created this defense in the form of UV screening compounds such as scytonemin. Scytonemin **(24)** is an extracellular sheath pigment first observed in 1849 when Nägeli described a yellow-green pigmentation in the sheaths of cyanobacteria (Garcia-Pichel and Castenholz, 1991). In 1993, the elusive unique dimeric indole-phenolic structure was elucidated (Proteau *et al.*, 1993). Scytonemin is produced solely by cyanobacteria and is considered a true sunscreen (Cockell and Knowland, 1999). This pigment absorbs strongly in the UV-A region with an *in vivo* λ_{max} = 370 nm and is found in at

least 30 species of sheathed cyanobacteria from various geographic locations and environments. Scytonemin effectively screens UV radiation by decreasing 85-90% of the incident UV-A radiation from reaching the interior of the cell (Proteau *et al.*, 1993).



(24) Scytonemin

Figure 1.11: Chemical structure of scytonemin

In the modern environment, cyanobacteria are ecologically successful as witnessed through their abundance in many habitats throughout the planet including those exposed to high levels of light. The current time period of global climate change requires an understanding of the effects of changing physical factors on biological systems (Paul, 2008). Part of global climate change involves stratospheric-ozone depletion leading to increases in UV radiation reaching the surface of the planet and an increasing concentration of atmospheric CO_2 contributing to sea-level rise and alterations in ocean chemistry. Interestingly, historic analyses indicate that in past periods of faunal mass extinction which coincided with global climate change,

cyanobacterial communities have both changed and increased in abundance. Cyanobacteria are even considered a “disaster species” in terms of being major components of coral reef communities during times of reef collapse. It seems that cyanobacteria may be more productive due to the environmental changes associated with global warming (Paul, 2008). Understanding the biological capabilities and adaptations that cyanobacteria have developed for these environmental changes may allow us to better harness them as sources of renewable resources in the future, as well as predict the microbial population changes that will occur as a result of global climate change.

GENERAL THESIS CONTENTS

The search for natural products from cyanobacteria with pharmacological and biotechnological potential has been the main focus of our laboratory for many years. This search has extended into the novel biosynthetic pathways responsible for the production of these natural products and a curiosity about the regulation of these unique gene clusters. An example of a cyanobacterial natural product under stringent regulation by an environmental parameter is scytonemin. Scytonemin's unusual molecular skeleton, unique biosynthetic pathway, and ecological importance as a UV-absorbing sunscreen in cyanobacteria are the unifying themes outlined in this thesis.

Following the general introduction to natural products and the evolutionary history of cyanobacteria in relation to UV radiation, the second chapter describes the bioinformatic exploration of *Nostoc punctiforme* L29133 for clues to the biosynthetic genes involved in the production of scytonemin. Genetic regions of interest were identified based on sequence similarity to a suite of genes involved in aromatic amino acid biosynthesis and metabolism, signaling, transport, and oxidative coupling hypothesized to play a role in scytonemin's biosynthesis. Genes from these regions were examined by semi-quantitative RT-PCR to evaluate for differential transcription after induction with UV radiation, and thus, to identify genetic regions responsible for the biosynthesis of scytonemin. One of these genetic regions was later shown by a lab at the University of Arizona using transposon mutagenesis to be responsible for scytonemin biosynthesis.

Chapter three discusses the scytonemin gene cluster in terms of its evolutionary history and its stringent regulation. The genes present in the cluster and those of the surrounding area were studied for their differential transcription after exposure to UV radiation revealing defined cluster boundaries. This cluster is also compared among genomes of various cyanobacterial taxa to identify interesting differences in the genetic architecture of this gene cluster and to establish its ancient evolutionary history through phylogenetic analysis.

The fourth chapter presents the biogenetic origins of the carbons involved in forming the scytonemin chemical structure through stable isotope incubation studies. These studies reveal tryptophan and tyrosine to be biosynthetic precursors and establish the *in vivo* carbon connectivities involved in the condensation of these precursors. This chapter also gains a unique insight into the induction and rate of scytonemin biosynthesis using MALDI-TOF mass spectrometry in the cyanobacterium *Tolypothrix* sp.

Chapter five focuses on the biochemical potential of the Np1263 gene product from the scytonemin gene cluster in *N. punctiforme* ATCC 29133. This study compares the reactivity of a recombinant protein derived from Np1263 with the well characterized mushroom tyrosinase. The results reveal the oxygenase function of Np1263 on a synthetic mimic of the proposed scytonemin monomer.

The thesis ends with a concluding chapter that contains a brief summary of the projects described within the text as well as comments on the importance of understanding scytonemin biosynthesis and trends in future research.

References

Albrizio, S., Ciminiello, P., Fattorusso, E., and Magno, S. (1995) Amphitoxin, a new high molecular weight antifeedant pyridinium salt from the Caribbean sponge *Amphimedon compressa*. *J. Nat. Prod*. 58: 647-652.

Bai, R., Pettit, G.R., and Hamel, E. (1990) Dolastatin 10, a powerful cytostatic peptide derived from a marine animal. *Biochem. Pharmacol*. 39: 1941-1949.

Bandaranayake, W.M. (1998) Mycosporines: are they nature's sunscreens? *Nat. Prod. Rep*. 15: 159-172.

Berman-Frank, I., Lundgren, P., and Falkowski, P. (2003) Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res. Microbiol*. 154: 157-164.

Blankenship, R.E. and Hartman, H. (1998) The origin and evolution of oxygenic photosynthesis. *Trends Biomed. Sci.* 23: 94-97.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., Currens, M.J., Cardellina, J.H., Buckheit, R.W., Nara, P.L., Pannell, L.K., Sowder, R.C., and Henderson, L.E. (1997) Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrob. Agents Chemother*. 41: 1521-1530.

Brielmann, H.L., Setzer, W.N., Kaufman, P.B., Kirakosyan, A., and Cseke, L.J. (1999) Phytochemicals: the chemical components of plants. In *Natural Products from Plants*, 2nd ed. CRC Press LLC, U.S.A. 1-36.

Burja, A.M., Banaigs, B., Abou-Mansour, E., Burgess, J.G., and Wright, P.C. (2001) Marine cyanobacteria – a prolific source of natural products. *Tetrahedron*. 57:9347-9377.

Butler, M.S. (2005) Natural products to drugs: natural product derived compounds in clinical trials. *Nat. Prod. Rep*. 22:162-195.

Carmichael, W.W., Azevedo, S.M.F.O., An, J.S., Molica, R.J.R., Jochimsen, E.M., Lau, S., Rinehart, K.L., Shaw, G.R., and Eaglesham, G.K. (2001) Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins. Environ. Health Perspect. 109: 663-668.

Carmichael, W. (2008) Chapter 4: a world overview – one-hundred-twenty-seven years of research on toxic cyanobacteria – where do we go from here? *Adv. Exp. Med. Biol*. 619: 105-125.

Castenolz, R.W. and Garcia-Pichel, F. (2000) Cyanobacterial responses to UV radiation. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M. Kluwer Academic Publishers: Netherlands, 591- 611.

Clark, A.M. (1996) Natural products as a resource for new drugs. *Pharm. Res*. 13: 1133-1141.

Cockell, C.S. and Knowland, J. (1999) Ultraviolet radiation screening compounds. *Biol. Rev*. 74: 311-345.

Cohen, T., and Gurevitz, M. (2006) The cyanobacteria – ecology, physiology and molecular genetics. *Prokaryotes*. 4: 1074-1098.

Cragg, G.M., Newman, D.J., and Snader, K.M. (1997) Natural products in drug discovery and development. *J. Nat. Prod*. 60: 52-60.

Crews, P., Rodriguez, J., and Jaspars, M. (1998) Optical and chiroptical techniques: ultraviolet spectroscopy. In *Organic Structure Analysis*. Oxford University Press: Oxford, 349-409.

D'Agostino, G., Del Campo, J., Mellado, B., Izquierdo, M.A., Minarik, T., Cirri, L., Marini, L., Perez-Garcia, J.L., and Scambia, G. (2006) A multicenter phase II study on the cryptophycin analog LY355703 in patients with platinum-resistant ovarian cancer. *Int. J. Gynecol. Cancer*. 16: 71-76.

Des Marais, D.J. (2000) When did photosynthesis emerge on Earth? *Science*. 289: 1703-1705.

Dismukes, G.C., Klimov, V.V., Baranov, S.V., Kozlov, Y.N., DasGupta, J., and Tyryshkin, A. (2001) The origin of atmospheric oxygen on Earth: the innovation of oxygenic photosynthesis. *Proc. Nat. Acad. Sci*. U.S.A. 98: 2170-2175.

Dittmann, E., Neilan, and Börner, T. (2001) Molecular biology of peptide and polyketide biosynthesis in cyanobacteria. *Appl. Microbiol. Biotechnol*. 57: 467-473.

Edge, R., McGarvey, D.J., and Truscott, T.G. (1997) The carotenoids as anti-oxidants – a review. *J. Photochem. Photobiol. B*. 41: 189-200.

Edwards, D.J., Marquez, B.L., Nogle, L.M., McPhail, K., Goeger, D.E., Roberts, M.A., and Gerwick, W.H. (2004) Structure and biosynthesis of the Jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium Lyngbya majuscule. *Chem. Biol*. 11: 817-833.

Ehling-Schulz, M. and Scherer, S. (1999) UV protection in cyanobacteria. *Eur. J. Phycol.* 34: 329-338.

Faulkner, D.J. (2000) Highlights of marine natural products chemistry (1972-1999). *Nat. Prod. Rep*. 17: 1-6.

Fischbach, M.A. and Walsh, C.T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev*. 106: 3468-3496.

Fleming, A. (1944) The discovery of penicillin. *Br. Med. Bull.* 2: 4-5.

Garcia-Pichel, F. and Castenholz, R.W. (1991) Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycol*. 27: 395-409.

Garcia-Pichel, F. (1998) Solar ultraviolet and the evolutionary history of cyanobacteria. *Origins Life Evol. Biosphere*. 28: 321-347.

Gerwick, W.H., Proteau, P.J., Nagle, D.G., Hamel, E., Blokhin, A., and Slate, D.L. (1994) Structure of curacin A, a novel antimitotic, antiproliferative and brine shrimp toxic natural product from marine cyanobacterium Lyngbya majuscule. *J. Org. Chem*. 59: 1243-1245.

Hirschberg, J. and Chamovitz, D. (1994) Carotenoids in cyanobacteria. In *The Molecular Biology of Cyanobacteria*. Bryant, D.A. Kluwer Academic Publishers, The Netherlands. 559-579.

Jaiswal, P., Singh, P.K., and Prasanna, R. (2008) Cyanobacterial bioactive molecules – an overview of their toxic properties. *Can. J. Microbiol*. 54: 701-717.

Jonasson, S., Eriksson, J., Berntzon, L., Rasmussen, U., and Bergman, B. (2008) A novel cyanobacterial toxin (BMAA) with potential neurodegenerative effects. *Plant Biotechnol*. 25: 227-232.

Kerksiek, K., Mejillano, M.R., Schwartz, R.E., Georg, G.I., and Himes, R.H. (1995) Interaction of cryptophycin 1 with tubulin and microtubules. *FEBS Lett*. 377: 59-61.

Kruschel, C. and Castenholz, R.W. (1998) The effect of solar UV and visible irradiance on the vertical movements of cyanobacteria in microbial mats of hypersaline waters. *FEMS Microbiol. Ecol*. 27: 53-72.

Kubanek, J., Whalen, K.E., Engel, S., Kelly, S.R., Henkel, T.P., Fenical, W., Pawlik, J.R. (2002) Multiple defensive roles for triterpene glycosides from two Caribbean sponges. *Oecologia*. 131: 125-136.

LePage, K.T., Goeger, D., Yokokawa, F., Asano, T., Shioiri, T., Gerwick, W.H., and Murray, T.F. (2005) The neurotoxic lipopeptide kalkitoxin interacts with voltage-sensitive sodium channels in cerebellar granule neurons. *Toxicol. Lett*. 158: 133-139.

Li, W.I., Berman, F.W., Okino, T., Yokokawa, F., Shioiri, T., Gerwick, W.H., and Murray, T.F. (2001) Antillatoxin is a marine cyanobacterial toxin that potently activates voltage-gated sodium channels. *Proc. Nat. Acad. Sci*. U.S.A. 98: 7599-7604.

Luesch, H., Moore, R.E., Paul, V.J., Mooberry, S.L., and Corbett, T.H. (2001) Isolation of dolastatin 10 from the marine cyanobacterium Symploca species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1. *J. Nat. Prod*. 64: 907-910.

Marquez, B.L., Watts, K.S., Yokochi, A., Roberts, M.A., Verdier-Pinard, P., Jimenez, J.I., Hamel, E., Scheuer, P.J., and Gerwick, W.H. (2002) Structure and absolute stereochemistry of hectochlorin, a potent stimulator of actin assembly. *J. Nat. Prod*. 65: 866-871.

Martins, J.C. and Vasconcelos, V.M. (2009) Microcystin dynamics in aquatic organisms. *J. Toxicol. Environ. Health, Part B*. 12: 65-82.

Mathews, C.K., van Holde, K.E., and Ahern, K.G. (2000) Photosynthesis. In *Biochemistry* 3rd edition. Addison Wesley Longman, Inc: U.S.A., 601.

Namikoshi, M. and Rinehart, K.L. (1996) Bioactive compounds produced by cyanobacteria. *J. Ind. Microbiol.* 17: 373-384.

Newman, D.J. and Cragg, G.M. (2007) Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod*. 70: 461-477.

Olson, J.M. and Pierson, B.K. (1986) Photosynthesis 3.5 thousand million years ago. *Photosynth. Res*. 9: 251-259.

Parson, E.A. (2003) Early stratospheric science, chlorofluorocarbons, and the emergence of environmental concerns. In *Protecting the Ozone Layer, Science and Strategy*. Oxford University Press. U.S.A. 14-25.

Paul, V.J. (2008) Global warming and cyanobacterial harmful algal blooms. In *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*. Springer. New York. 619: 239-257.

Pettit, G.R., Kamano, Y., Herald, C.L., Tuinman, A.A., Boettner, F.E., Kizu, H., Schmidt, J.M., Baczynskyj, L., Tomer, K.B., and Bontems, R.J. (1987) The isolation and structure of a remarkable marine animal antineoplastic constituent: dolastatin 10. *J. Am. Chem. Soc*. 109: 6883-6885.

Proteau, P.J., Gerwick, W.H., Garcia-Pichel, F., and Castenholz, R. (1993) The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experentia*. 49: 825-829.

Rowland, F.S. (2006) Stratospheric ozone depletion. *Philos. Trans. R. Soc. B*. 361: 769-790.

Schopf, J.W. (1994) Disparate rates, differing fates: Tempo and mode of evolution changed from the Precambrian to the Phanerozoic. *Proc. Nat. Acad. Sci*. U.S.A. 91: 6735-6742.

Simmons, T.L., Andrianasolo, E., McPhail, K., Flatt, P., and Gerwick, W.H. (2005) Marine natural products as anticancer drugs. *Mol. Cancer Ther*. 4: 333-342.

Sinha, R.P., Singh, N., Kumar, A., Kumar, H.D., Häder, M., and Häder, D.-P. (1996) Effects of UV irradiation on certain physiological and biochemical processes in cyanobacteria. J. Phytochem. Photobiol. B: Biology. 32: 107-113.

Smith, C.D., Zhang, X., Mooberry, S.L., Patterson, G.M.L., and Moore, R.E. (1994) Cryptophycin: a new antimicrotubule agent active against drug-resistant cells. *Cancer Res*. 54: 3779-3784.

Snyder, L. and Champness, W. (2003) DNA repair and mutagenesis. In *Molecular Genetics of Bacteria*, 2nd ed. ASM Press: U.S.A., 385-387.

Stal, L.J. (2000) Cyanobacterial mats and stromatolites. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M., eds. Kluwer Academic Publishers: Netherlands, 61-120.

Summons, R.E., Jahnke, L.L., Hope, J.M., and Logan, G.A. (1999) 2-methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature*. 400: 554-557.

Takaichi, S. and Mochimaru, M. (2007) Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides. *Cell. Mol. Life Sci*. 64: 2607-2619.

Tan, L.T. (2007) Bioactive natural products from marine cyanobacteria for drug discovery. *Phytochem*. 68: 954-979.

Tidgewell, K., Clark, B.R., and Gerwick, W.H. (2009) The natural products chemistry of cyanobacteria. In *Comprehensive Natural Product Chemistry*, 2nd ed. Moore, B. and Crews, P. eds. Elsevier: Oxford (In Press).

Tomitani, A., Knoll, A.H., Cavanaugh, C.M., and Ohno, T. (2006) The evolutionary diversification of cyanobacteria: molecular-phylogenetic and palenotological perspectives. *Proc. Nat. Acad. Sci*. U.S.A. 103: 5442-5447.

Whitton, B.A., and Potts, M. (2000) Introduction to the cyanobacteria. In *The Ecology of Cyanobacteria*. Kluwer Academic Publishers. Netherlands. 1-11.

Wipf, P., Reeves, J.T., Balachandran, R., and Day, B.W. (2002) Synthesis and biological evaluation of structurally highly modified analogues of the antimitotic natural product curacin A. J. *Med. Chem*. 45: 1901-1917.

Woese, C.R. (1987) Bacterial evolution. *Microbiol. Mol. Biol. Rev*. 51: 221-271.

Wynn-Williams, D.D., Edwards, H.G.M., Newton, E.M., and Holder, J.M. (2002) Pigmentation as a survival strategy for ancient and modern photosynthetic microbes under high ultraviolet stress on planetary surfaces. *Int. J. Astrobiol*. 1: 39-49.

CHAPTER TWO -

EXPLORATION OF THE *NOSTOC PUNCTIFORME* ATCC 29133 GENOME FOR CLUES TO UNDERSTANDING SCYTONEMIN BIOSYNTHESIS

Abstract

Scytonemin is a yellow to brown pigment found in the sheaths of many species of cyanobacteria. This pigment's unique dimeric indole-phenolic molecular structure protects the cyanobacterial cells from exposure to UV radiation. Scytonemin's highly conjugated structure is predicted to be biosynthesized from the condensation of tryptophan and tyrosine derived subunits. In this study, we use insights from this chemical structure to identify candidate genes from *Nostoc punctiforme* ATCC 29133 involved in the biosynthesis of scytonemin through bioinformatic analyses. We also show the induced upregulation of three of the four candidate gene regions by UV radiation.

Introduction

Scytonemin is an extracellular sheath pigment first described as a yellow to brown coloration in cyanobacteria by Nägeli in 1849. This pigment is known to be produced by over 30 species of cyanobacteria (Garcia-Pichel and Castenholz, 1991). Scytonemin's highly conjugated and unique dimeric indole-phenolic chemical structure absorbs strongly in the UV-A range (315-400 nm) with a λ_{max} = 370 nm *in*

vivo. It also has significant absorption in both the UV-B (280-320 nm) and UV-C (200-280 nm) ranges (Proteau *et al*., 1993). The production of scytonemin is elicited by the exposure of cyanobacteria to UV radiation, either in the field or laboratory culture. The combination of UV radiation with other environmental stress factors including temperature, osmotic and oxidative stress can significantly increase the level of scytonemin (Dillon *et al*., 2002). Scytonemin's role as a protective screening pigment in cyanobacteria is supported by its ability to screen 95% of the UV-A photons from entering the cells based on decreases in red auto-flourescence induced by UV-A radiation in cells without scytonemin (Garcia-Pichel, 1991; Garcia-Pichel *et al*., 1992). The screening capacity of this pigment is known to persist in cyanobacterial cultures within empty sheaths and during times of cyanobacterial physiological inactivity (Dillon and Castenholz, 1999).

Scytonemin's unique ring system, termed the "scytoneman" skeleton, was first elucidated in 1993 and predicted to be derived from the condensation of aromatic amino acid derived subunits (Proteau *et al*, 1993). The aromatic amino acids (AAAs) consist of tyrosine (Tyr), tryptophan (Trp), and phenylalanine (Phe). These amino acids are categorized based on the presence of an aromatic ring structure as a side group, and constitute less than 10% of the total number of amino acids in proteins on average due to their high biosynthetic cost (Sprenger, 2006). The biosynthesis of AAAs occurs via a conserved biosynthetic pathway in microorganisms, fungi and plants, but not in animals, who usually require supplementation of these amino acids (Herrmann and Weaver, 1999). The AAA biosynthetic pathway consumes two

molecules of phosphoenolpyruvate and one molecule of erythrose-4-phosphate to form chorismate, the precursor to all three AAAs. Following the formation of chorismate, the pathway splits into two branches. One branch is responsible for tyrosine and phenylalanine biosynthesis via the intermediate prephenic acid and the other branch is responsible for tryptophan biosynthesis via shikimic acid (Sprenger, 2006). An exception to this common pathway is found in some microorganisms including cyanobacteria where the biosynthesis of phenylalanine and tyrosine occurs through the intermediate L-arogenate (Hall, 1982). This cyanobacterial biosynthetic pathway is outlined in **figure 2.1**.

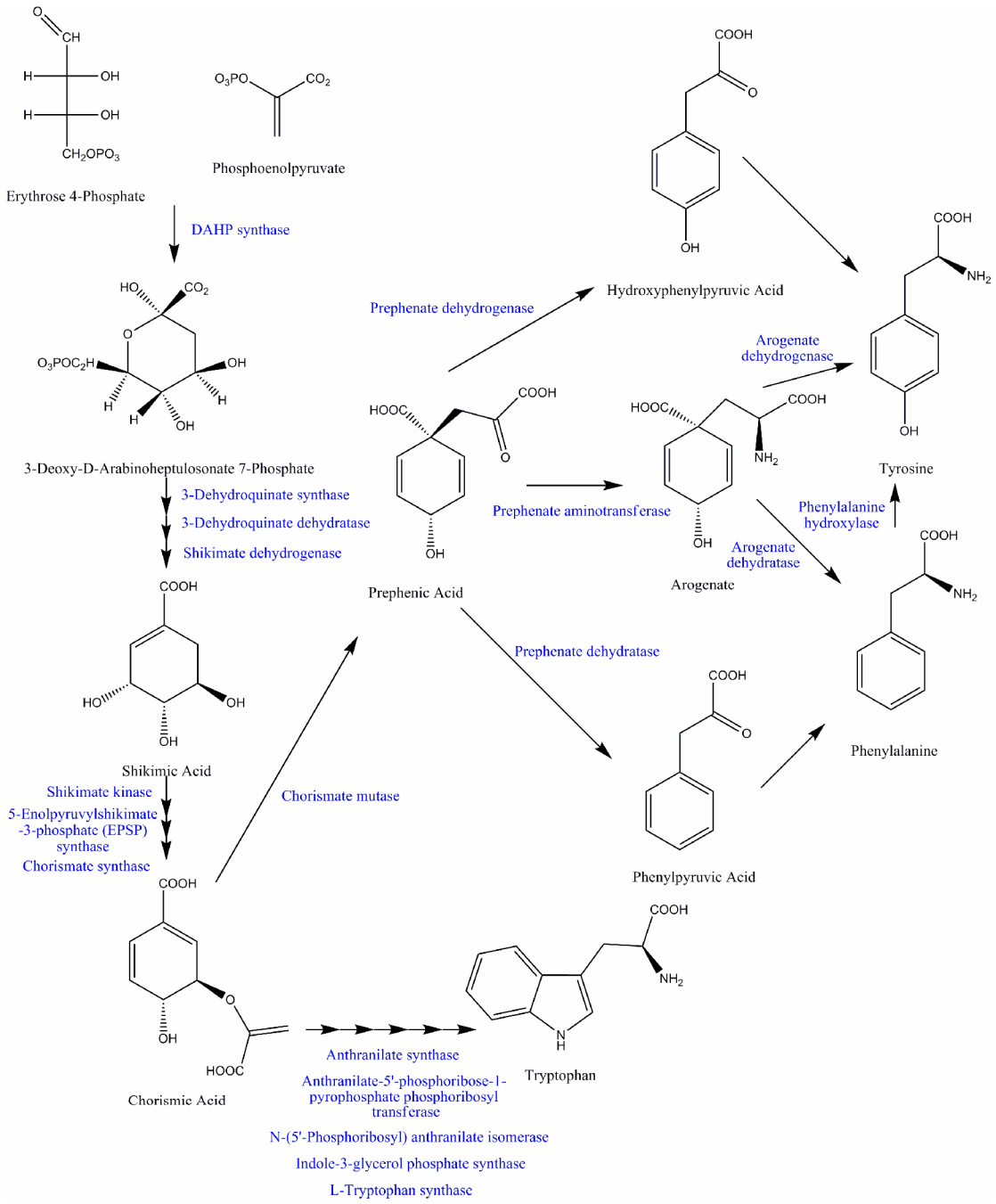


Figure 2.1: Diagram of the common aromatic amino acid biosynthetic pathway in microorganisms. Enzymes involved in biosynthesis are shown in blue (Hall, *et al.*, 1982).

The ability of AAAs to absorb UV radiation at 260 nm (Phe) and 280 nm (Tyr and Trp) make them susceptible to the damaging effects of these wavelengths, particularly UV-C (Sprenger, 2006). However, the condensation of these simple molecules can form more complex molecules that absorb at higher wavelengths of UV radiation due to increasing conjugation. These larger wavelengths, particularly UV-A, are important contributors to biological damage in the modern environment (Cockell, 1999). AAAs play a significant role in many of the metabolites used as a response to UV radiation across both the eukaryotic and prokaryotic kingdoms, including scytonemin in cyanobacteria, phenylpropanoids in plants and melanins in animals (Knaggs, 2003).

The combination of tryptophan and tyrosine derived subunits to form scytonemin's highly conjugated structure led to an interest in its unusual biosynthesis. *Nostoc punctiforme* ATCC 29133 is a nitrogen fixing cyanobacterium with a wide range of vegetative cell developmental patterns and ecological niches. This cyanobacterium was previously shown to produce scytonemin upon exposure to UV radiation and has a published genome (Meeks *et al*., 2001; Hunsucker *et al*., 2004). The focus of this study was to use insights from the predicted mechanism for scytonemin biosynthesis, largely based on its unique structure, to identify genes involved in scytonemin biosynthesis. This investigation used bioinformatic analyses of the *N. punctiforme* ATCC 29133 genome combined with differential expression

techniques including suppressive subtractive hybridization and semi-quantitative reverse transcription PCR (sRT-PCR) to identify genes potentially involved in scytonemin biosynthesis.

Materials and Methods

Cyanobacterial Strains and Culture Techniques

The cyanobacterium *Nostoc punctiforme* ATCC 29133 was obtained from the American Type Culture Collection (ATCC). A culture was maintained in unialgal condition in Allen and Arnon (AA) freshwater media at 29^{o}C under a light intensity of approximately 19 μmol m^{-2} s^{-1} and a light/dark cycle of 16 h/8 h.

Bioinformatic Analyses

The genome sequence for *N. punctiforme* ATCC 29133 (GenBank #CP001037) was obtained from the website for the DOE Joint Genome Institute (JGI) with genome annotations completed by the Computational Biology group at the Oak Ridge National Laboratory (ORNL). Nucleotide and amino acid sequences for genes with functions predicted to be involved in scytonemin biosynthesis or involving aromatic amino acids were identified using the National Center for Biotechnology Information (NCBI) and BRENDA (Department of Bioinformatics and Biochemistry at Technische Universität Carolo-Wilhelmina zu Braunschweig). The accession numbers for genes used in this analysis are included in **table 2.2** at the end of the chapter. These amino acid sequences were used with the Basic Local Alignment Search Tool (BLAST, NCBI) against the *N. punctiforme* ATCC 29133 genome. The

genome annotations were also used to identify candidate genes. The resulting candidate genes found in *N. punctiforme* ATCC 29133 were then located on the genome to find genetic regions containing multiple genes with functions predicted to be involved in scytonemin biosynthesis. These functions included a relationship to aromatic amino acids, signaling elements, transport capabilities, and sugar transfer or cleavage capabilities.

Suppressive Subtractive Hybridization

N. punctiforme ATCC 29133 was grown for approximately 75 days under the light and growth conditions described above and then exposed to UV-A radiation (λ_{max} = 365 nm) for 48 h in an open pan. Cultured biomass was collected from the pan prior to UV radiation induction and after 48 h for extraction of RNA following a modified Trizol protocol (Invitrogen). The extracted RNA was treated with Turbo DNase (Ambion). RNA concentrations were determined using a Beckman Coulter DU800 spectrophotometer. The integrity and purity of RNA was also analyzed using a denaturing-formaldehyde gel and an Agilent 2100 bioanalyzer (Biogem, UCSD). Complementary DNA (cDNA) was created through a series of three amplification steps using random primers. The random primers included R1 (CTGCTTGATGAAA), R2 (CTGCTTGATGAAC), R3 (CTGCTTGATGAAG), Lnr81 (TGAGCGGACA), Lnr95 (CAGCCCAGAG), and Lnr99 (TCGTGCGGGT; Frias-Lopez *et al*., 2004). The first strand cDNA was created using 2 µg of total RNA, 1 µL 10mM dNTPs, 1 µL random primer, 4 µL 5X first strand reaction buffer, 1 µL 0.1M DTT, 1 µL RNaseOut (Invitrogen), 0.5 µL Superscript III reverse transcriptase

(Invitrogen), and 8 µL sterile water per reaction. Each reaction was incubated at 4°C for 2 minutes followed by 37°C for 60 minutes. Prior to the reaction, negative controls were incubated at 95°C for 5 minutes to denature the enzyme. The first strand cDNA was then used in a low stringency PCR reaction consisting of 2 µL first strand template, 2 µL 10X PCR buffer, 0.6 µL 50mM $MgCl_2$, 0.4 µL 10mM dNTPs, 1 µL random primer, 0.4 µL Platinum Taq polymerase (Invitrogen), and 15.2 µL sterile water per reaction. The low stringency reaction conditions were 94°C for 3 min followed by 15 cycles of 94°C for 1 min, 37°C for 1 min, 72°C for 2 min, and completed with a five minute incubation at 72°C. The low stringency reaction product was then used as a template in a high stringency PCR reaction consisting of 2 µL template, 2 µL random primers, 5 µL 10x reaction buffer, 1 µL 10mM dNTPs, 0.5 µL Taq polymerase (Invitrogen), and 37.5 µL sterile water per reaction. These reactions followed the same conditions as the low stringency except with an annealing temperature of 50°C.

Once cDNA was created using all six random primers, they were combined and purified using isopropanol repurification. Both –UV and +UV cDNA were digested with RsaI for 6 hours following a standard protocol (Invitrogen). The –UV cDNA was purified and stored at -20°C while +UV cDNA was split into two aliquots. Each +UV aliquot was ligated with one of two adaptors, Adaptor 1 (5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3') or Adaptor 2R (5'-CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGA

GGT-3') using T4 DNA ligase and standard conditions (New England Biolabs). Following ligation, the efficiency was determined through PCR using primers designed to bridge the ligation site between the adaptor and a specific gene (Np4631).

Following confirmation of adaptor ligation, the +UV adaptor ligated cDNA was incubated with the digested –UV cDNA in hybridization buffer (50mM HEPES pH 8.3, 0.5 M NaCl, 0.02 mM EDTA pH 8.0, 10% w/v PEG-8000) overnight at 68^{o}C (Clontech PT1117-1). The cDNA with each of the adaptors was incubated separately with –UV cDNA. Following the first hybridization, the hybridized cDNA for both adaptors were combined with an excess of undenatured –UV RsaI digested cDNA and incubated overnight at 68^{o}C. The series of hybridizations was followed by two PCR reactions. The first PCR reaction was designed to fill in the gaps at the end of the products. This reaction consisted of 1 µL subtracted cDNA, 2.5 µL 10X PCR buffer, 0.5 µL 10mM dNTPs, 1 µL PCR Primer 1 (5'-CTAATACGACTCACTATAGGGC-3'), 0.5 µL 50X Advantage Polymerase mix (Clontech), and 19.5 µL sterile water per reaction. The reaction conditions were 75^{o}C for 5 min, 94^{o}C for 2 min, 27 cycles of 94^{o}C 30 sec, 66^{o}C 30 sec, 72^{o}C 1.5 min, followed by a final incubation of 72^{o}C for 5 min. The second PCR reaction designed to amplify these products consisted of 1 µL of template from the first reaction, 2.5 µL 10X PCR buffer, 1 µL nested primer 1 (5'-TCGAGCGGCCGCCCGGGCAGGT-3'), 1 µL nested primer 2 (5'-AGCGTGGTCGCGGCCGAGGT-3'), 1 µL 10mM dNTPs, 0.5 µL Taq polymerase (Invitrogen), and 18 µL sterile water. The reaction conditions were 94^{o}C for 5 min, 25 cycles of 94^{o}C for 30 sec, a gradient of 55-68^{o}C for 30 sec, 72^{o}C for 1.5 min followed

by a final incubation of 72°C for 5 min. These PCR products were cloned into a TOPO pCR®4 cloning vector (Invitrogen) and transformed into Top10 *Escherichia coli* cells. After a standard PCR check for the presence of an insert, 22 clones were sent for sequencing (Sexcel, La Jolla, CA).

Differential Transcriptional Expression Analyses

N. punctiforme ATCC 29133 was grown for approximately 75 days under the light and growth conditions described above and then exposed to UV-A radiation (λ_{max} = 365nm) for 48 h in an open pan. Cultured biomass was collected from the pan prior to UV radiation induction and after 48 h for extraction of RNA following a modified Trizol protocol (Invitrogen). The extracted RNA was treated with Turbo DNase (Ambion). RNA concentrations were determined using a Thermo Scientific Nanodrop ND1000 spectrophotometer. Specific primers were designed to amplify regions of RNA for a specific gene within each gene region. These primer sets include NpF0025 (5'-ATGACAACCTCACAGGAG-3' and 5'-CATACTCGATATCGTGCAG-3'), NpF2914 (5'-GCGCTGACTATGGTGCTC-3' and 5'-TAGACCAACATGGGCATC-3'). cDNA was synthesized following a standard protocol for Superscript III reverse transcriptase with an incubation at 50°C (Invitrogen) and using 2.5 µg of RNA as template for every reaction. The PCR amplification was accomplished using a 0.5 µL Taq DNA polymerase (Invitrogen), 5 µL 10X PCR reaction buffer, 4 µL 10mM $MgCl_2$, 1 µL 10mM dNTPs, 2 µL of total primer, 1 µL of first strand cDNA template, and 37 µL sterile water per reaction. The PCR products were visualized on a 0.8%

agarose gel stained with ethidium bromide and documented using a Fisher Biotech transilluminator and a Kodak photographic system.

Phylogenetic Analyses

Phylogenetic analyses were completed using MEGA 4.0 (Tamura *et al*., 2007). All sequence alignments were performed using ClustalW aligorithms with the Gonnet protein weight matrix for amino acids. Neighbor-joining, minimum evolution, and maximum parsimony trees were created and evaluated with 10,000 bootstrap replicates. The outgroup used in the creation of the phylogenetic tree was obtained by identifying homologues by BLAST searches.

Results

Identifying Candidate Biosynthetic Genes

The predicted mechanism for scytonemin biosynthesis is based on its aromatic amino acid derived chemical structure. Potential biosynthetic routes for its production are complemented by the known dependence of cyanobacteria on UV radiation for the production of this molecule, indicating that signaling elements play a role in its regulation (Sinha and Häder, 2008). The predicted mechanism (**Figure 2.9**) was used to determine potential gene candidates involved in the biosynthesis of this pigment in *N. punctiforme* ATCC 29133. The *N. punctiforme* ATCC 29133 genome sequence is a single circular chromosome composed of 9,059,191 bp, with 5 plasmids, and an average G+C content of 41.4%. This genome is predicted to contain 6,690 protein coding genes (DOE JGI, 2009).

Potential gene candidates were discovered through BLAST searches using genes from other organisms with predicted functions related to scytonemin biosynthesis as a query. The genes identified from other organisms included genes involved in aromatic amino acid biosynthesis and metabolism (**Figure 2.1**), signaling elements such as histidine kinases and response regulators, transporters, genes involved in oxidative catalysis such as P450 enzymes, and genes involved in sugar transfer and cleavage such as glycosyltransferases and β-glucosidases (**Figure 2.9**). Accession numbers and host organisms for examples of the genes are listed in **table 2.2** at the end of this chapter. A total of 657 candidate genes with at least 30% sequence similarity to queried genes were identified out of the 6,798 annotated genes in *N. punctiforme* ATCC 29133 (Meeks *et al*., 2001). The number of candidate genes identified for each gene function category is summarized in **table 2.1**.

Table 2.1: Summary of the gene functions used for sequence similarity bioinformatics within the *Nostoc punctiforme* L29133 genome to identify candidate genes involved in scytonemin biosynthesis. The number of genes identified from the genome classified by category are also listed.

Gene Function Category	**Number of Genes with BLAST Similarity**
Two-Component Signaling	180
Other Signaling	25
Tyrosine Metabolism	29
Tryptophan Metabolism	13
Tryptophan Biosynthesis	46
Activation	58
Transporters	229
Glucosidases	8
Sugar Transfer	49
Oxidative Coupling	20
Total	**657**

The 657 candidate genes were categorized based on general functional classes as listed above to determine regions in the genome with higher concentrations of these candidate genes. Seventy-two genes located in regions containing these higher concentrations or having unique predicted functions related to aromatic amino acids such as transporters or decarboxylases were chosen for further bioinformatic analyses including gene neighborhood surveys. Based on the surrounding genetic region of these 72 candidate genes, four genetic regions of interest in the *N. punctiforme* ATCC 29133 genome were identified as potential candidates for involvement in scytonemin biosynthesis. The locations of these genetic regions in relation to the 657 original candidate genes are diagrammed in **figure 2.2**. These regions included those surrounding NpR0978, NpR1263, NpR1802, and NpR6003 as diagrammed in **figure 2.3**.

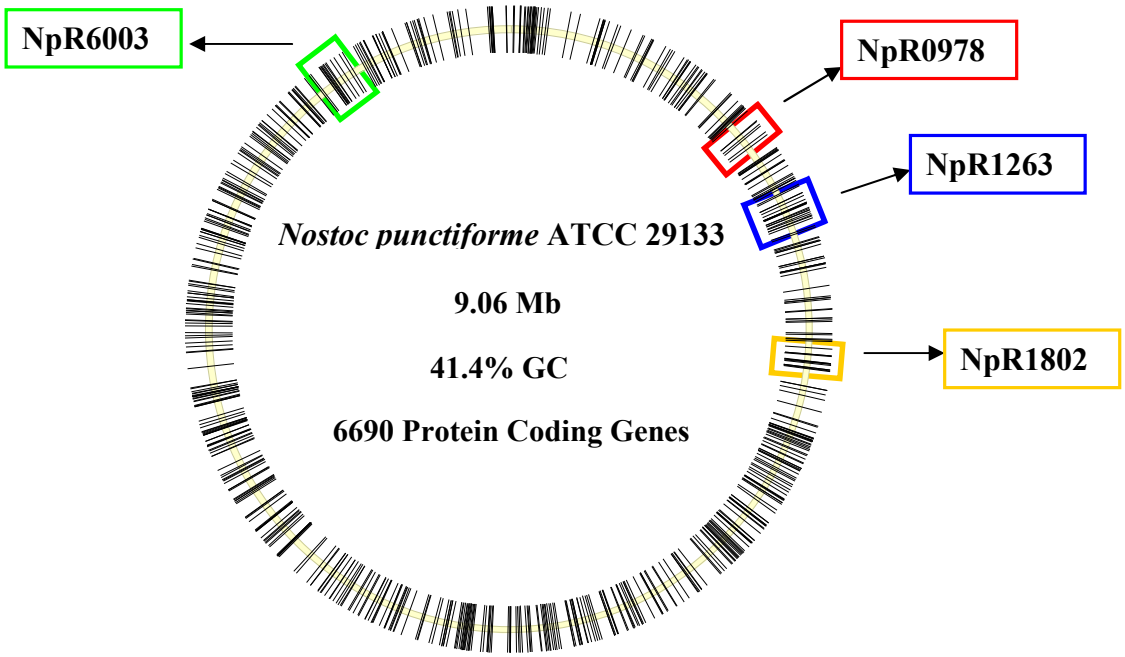


Figure 2.2: Diagram of the *N. punctiforme* ATCC 29133 genome (recreated from Meeks *et al*., 2001). Each black line represents one of the 657 candidate genes identified through BLAST. The colored boxes indicate the four candidate regions potentially involved in scytonemin biosynthesis. These regions are detailed in figure 2.3.

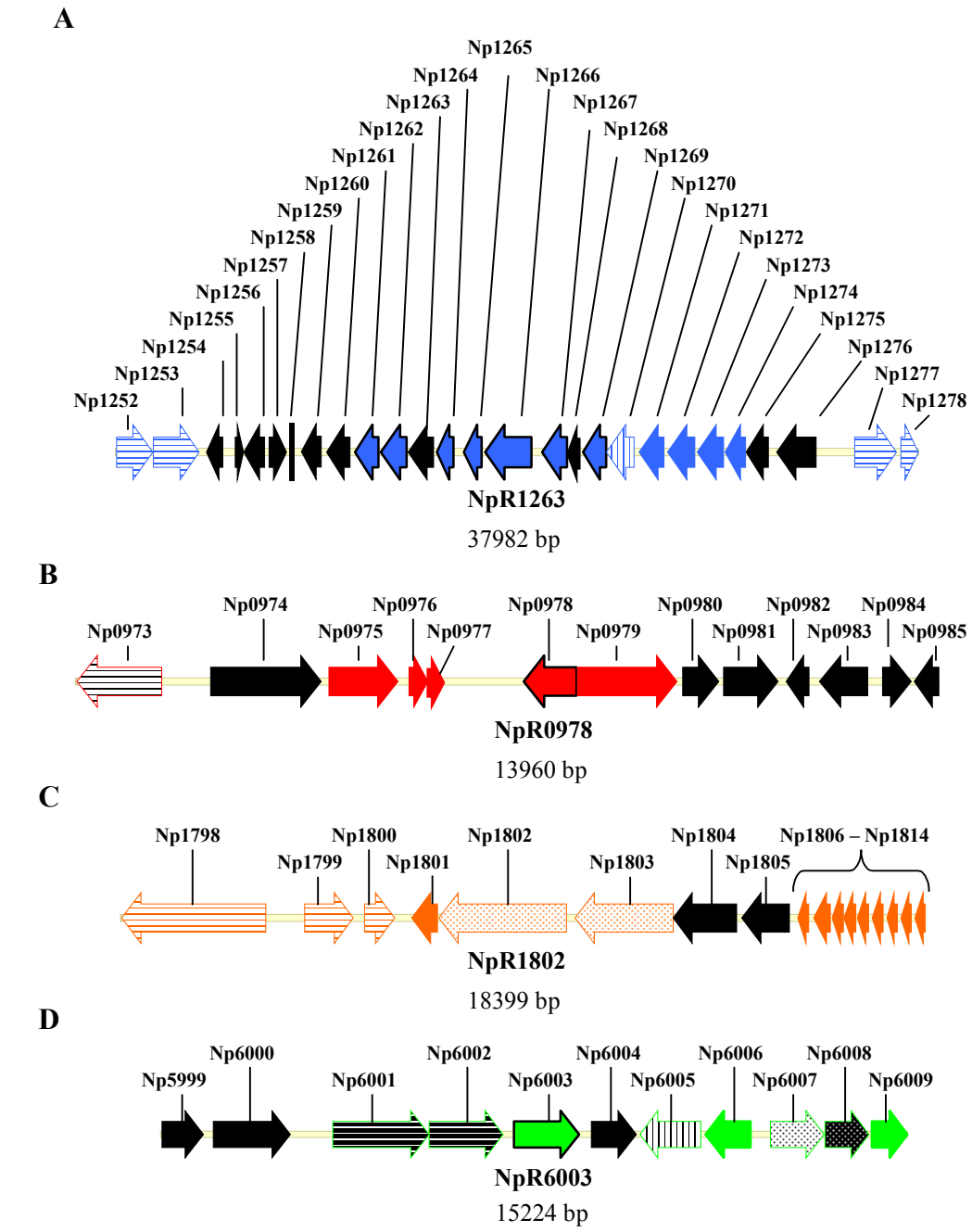


Figure 2.3: Diagram of the four candidate regions from figure 2.2. Arrows and lines represent annotated genes. Each gene is color coded based on predicted functional relationship to scytonemin biosynthesis. A) NpR1263 region, B) NpR0978 region, C) NpR1802 region, D) NpR6003 region.

Suppressive Subtractive Hybridization

SSH is a PCR based subtraction method that can selectively amplify differentially expressed cDNAs obtained from RNA. Total RNA isolated from both a control culture (-UV) and a culture exposed to UV radiation (+UV) were used to create cDNA. Total RNA was analyzed for integrity using a formaldehyde-denaturing gel (**Figure 2.4A**). This gel shows only three bands in the region of expected size for the 16S and 23S ribosomal subunits, indicating that the RNA has not degraded. The total RNA was also tested for genomic DNA contamination using PCR (**Figure 2.4B**). RNA was used in first strand cDNA synthesis with and without the presence of reverse transcriptase followed by PCR amplification. Visualization of the products on an agarose gel shows PCR products for lanes 2 and 4 representing the cDNA created with reverse transcriptase, and no products for lanes 3 and 5 representing the cDNA created without reverse transcriptase. These results are consistent with RNA that is free of genomic DNA contamination. Once the quality of the RNA was tested, cDNA was amplified using reverse transcription and PCR. The quality of the cDNA was once again tested by PCR using primers designed for genes located throughout the genome (**Figure 2.4C**). The primers for Np0025, Np2914 and Np4104 all resulted in clean products at approximately 500bp. This result is consistent with the amplification of RNA from throughout the genome to create cDNA. Maintaining the variation in the transcriptome throughout the amplification of cDNA can be verified by visualizing the –UV and +UV cDNA fingerprint on an agarose gel (**Figure 2.4D**). The gel shows

an increased band intensity for the +UV cDNA at arrows A, B, E, and H, while the arrow for C, D, and F show a decreased intensity in this gel.

The ligation between the cDNA and the adaptors was verified based on PCR. This analysis resulted in a PCR product when amplification spanned the site of ligation between an adaptor and a primer designed from *N. punctif*orme ATCC 29133 for NpR4631. These PCR products were visualized on an agarose gel (**Figure 2.5**). Lanes 3 and 5 show the amplification of Np4631 from cDNA while Lanes 2 and 4 show the amplification of the adaptor ligated Np4631 based on the increased size of the PCR product. These results indicate that both adaptors 1 and 2R were ligated onto Np4631; however, the efficiency of ligation was low, particularly for adaptor 2R. The adaptor ligated cDNA was then used for the subtractive hybridization and ensuing PCR as outlined in **figure 2.6**. Subcloning and sequencing of the PCR products revealed the presence of a contaminant, *Sphingomonas* sp. KTO216, during the first experimental trial and resulted in only rRNA sequences from *N. punctiforme* ATCC 29133 during the second experimental trial. The resulting sequences were found to contain the adaptor sequence at both the 5'- and 3'- ends.

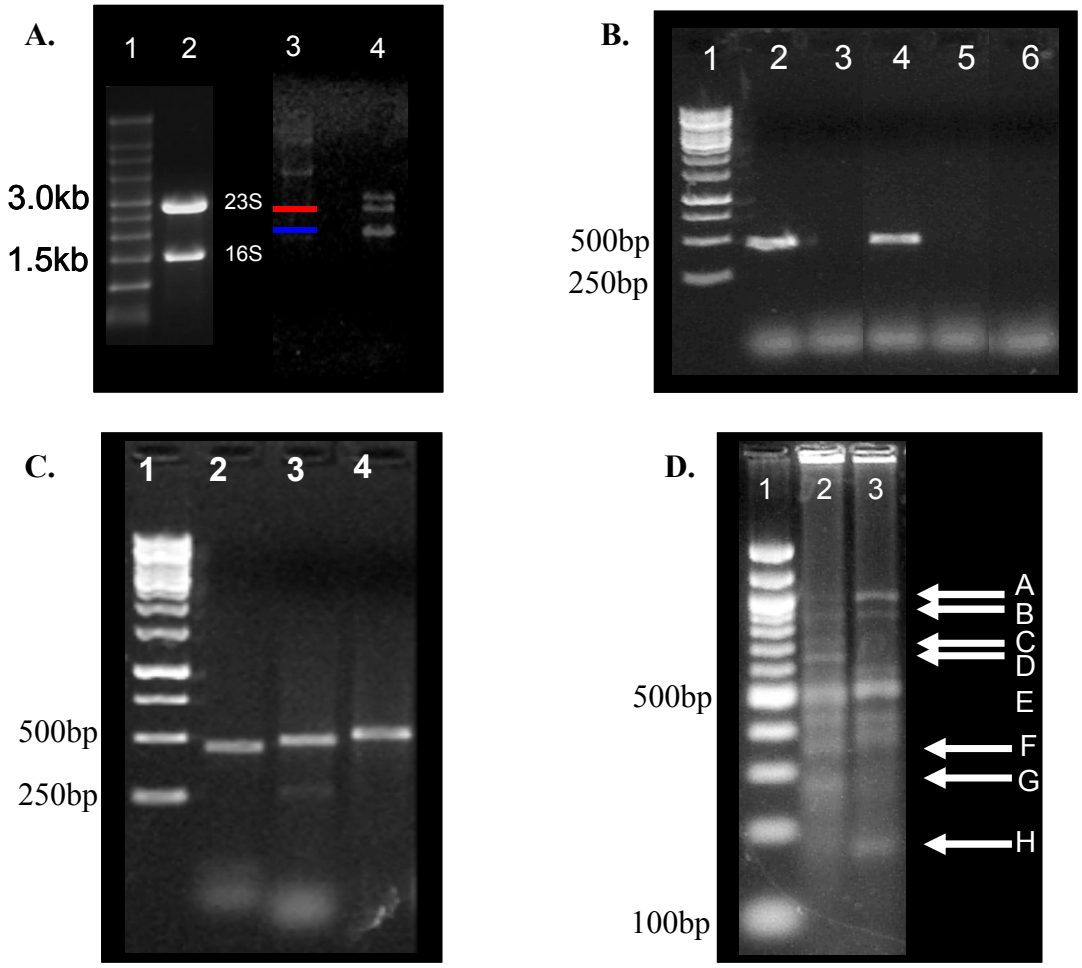


Figure 2.4: Visual analyses of RNA and cDNA quality during preparation for suppressive subtractive hybridization. A) Denaturing-formaldehyde gel stained with ethidum bromide (EtBr) showing RNA quality (1-Ambion Millinium Marker (Ambion), 2-*Pseudomonas aeruginosa* RNA (Ambion), 3-Invitrogen 9.5kb RNA Ladder, 4-*N. punctiforme* ATCC 29133 RNA). B) 0.8% Agarose gel stained with EtBr for analysis of genomic DNA contamination in RNA (1-1kb ladder, 2- -UV Np0025, 3- -UV Np0025 control, 4- +UV Np0025, 5- +UV Np0025 control, 6- Negative control). C) 0.8% Agarose gel stained with EtBr showing the quality of cDNA synthesis through the PCR analysis of genes scattered in the genome (1-1kb ladder, 2- Np0025, 3- Np2914, 4- Np4104). D) 0.8% Agarose gel stained with EtBr showing the cDNA fingerprint for –UV and +UV cDNA (1-100bp ladder, 2- -UV cDNA, 3- +UV cDNA).

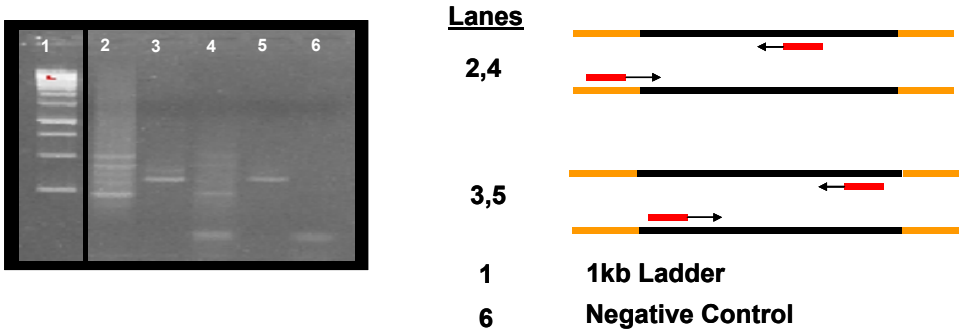


Figure 2.5: A 0.8% Agarose gel stained with EtBr showing the ligation efficiency of cDNA with adaptors 1 and 2R. Lanes 3 and 5 show primers within the cDNA and lanes 2 and 4 show primers designed to span the site of ligation as outlined on the right.

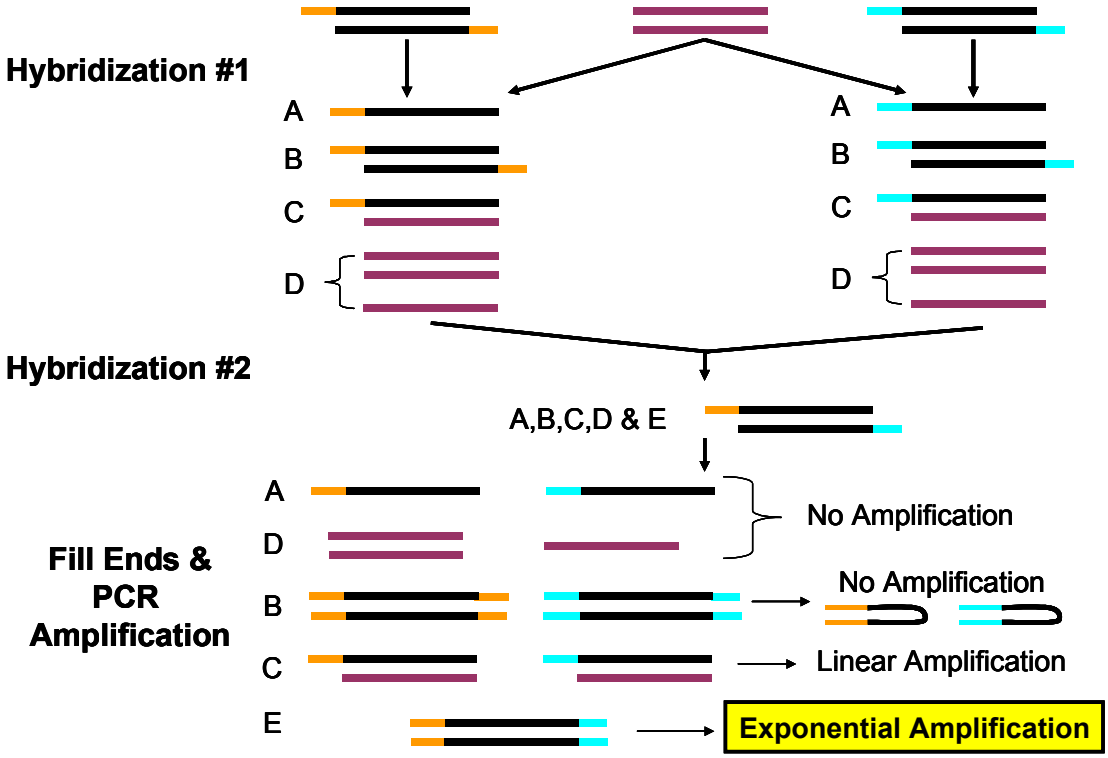


Figure 2.6: Diagram of steps used in suppressive subtractive hybridization including the predicted results from final amplification (Clontech).

Differential Transcription Analysis

Candidate genes from the four genetic regions of interest in *N. punctiforme* ATCC 29133 genome were analyzed for relative changes in their transcription after 24 and 48 hours of exposure to UV-A radiation. These transcript changes were visually compared to two controls, gyrase (Np0025) and RecA (Np2914). The visual inspection of the band intensity shows no apparent change between –UV and +UV transcription levels for gyrase and NpR1802; however, an increase in band intensity for the +UV products is present for RecA, NpR0978, NpR1263, and NpR6003 (**Figure 2.7**).

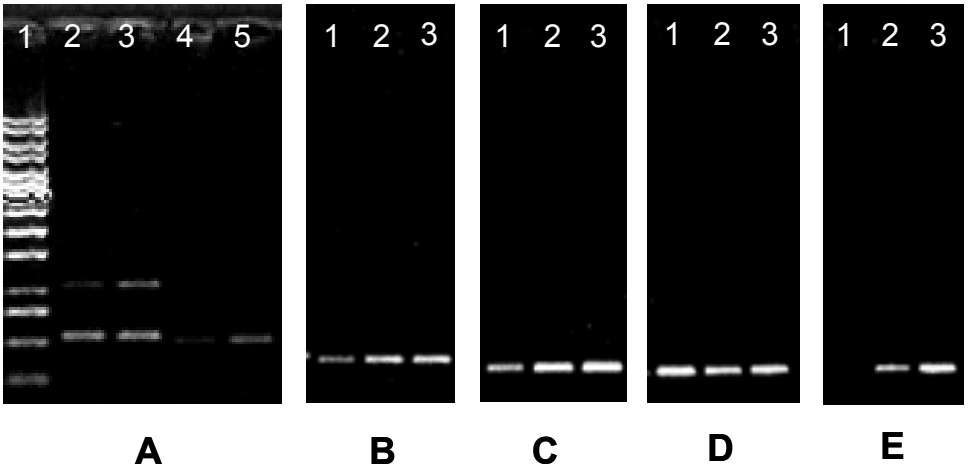


Figure 2.7: Agarose gel results of reverse transcription using specific primers designed from candidate genes in *N. punctiforme* ATCC 29133. In gels B-E, lanes are shown as follows: lane 1 - -UV, 2 - +UV 24hrs, 3 - +UV 48hrs.

A.) 1 – 1kb Ladder, 2 - -UV Gyrase, 3 - +UV Gyrase, 4 - -UV RecA, 5 - +UV RecA

B.) Gene Np0978 – Prephenate dehydrogenase

C.) Gene Np1263 -- Tyrosinase

D.) Gene Np1802 – ABC transporter

E.) Gene Np6003 – Putative tyrosine/tryptophan transport protein

Bioinformatic and Phylogenetic Analyses

Np1263 was identified from one of the most promising candidate regions for involvement in scytonemin biosynthesis. This gene was annotated as a hypothetical protein, but further BLAST analysis revealed significant sequence similarity (e-value $1e^{-31}$) to MelC2 from the melanin biosynthetic gene cluster in *Streptomyces avermitilis*.

The coupling of the scytonemin monomers to form scytonemin is predicted to be similar to the coupling of indican to form indigo. This coupling mechanism is driven by the cleavage of the sugar moiety by a β-glucosidase enzyme. The β-glucosidase amino acid sequence from *Polygonum tinctorium* was used in a tBLASTn search against the available cyanobacterial genomes. This search revealed five cyanobacterial species with significant sequence similarity including *Gloeobacter violaceus* PCC 7421, *Lyngbya* sp. PCC 8106, *Acaryochloris marina* MBIC11017, *Cyanothece* sp. PCC 7425, and *N. punctiforme* ATCC 29133. The gene from *N. punctiforme* ATCC 29133 (Np2285) is annotated as an extracellular solute binding protein in family 3. When other genes related to sugar cleavage mechanisms are compared with Np2285 and the β-glucosidase from *P. tinctorium*, Np2285 and *P. tinctorium* clade separately from these other sugar cleavage genes (**Figure 2.8**).

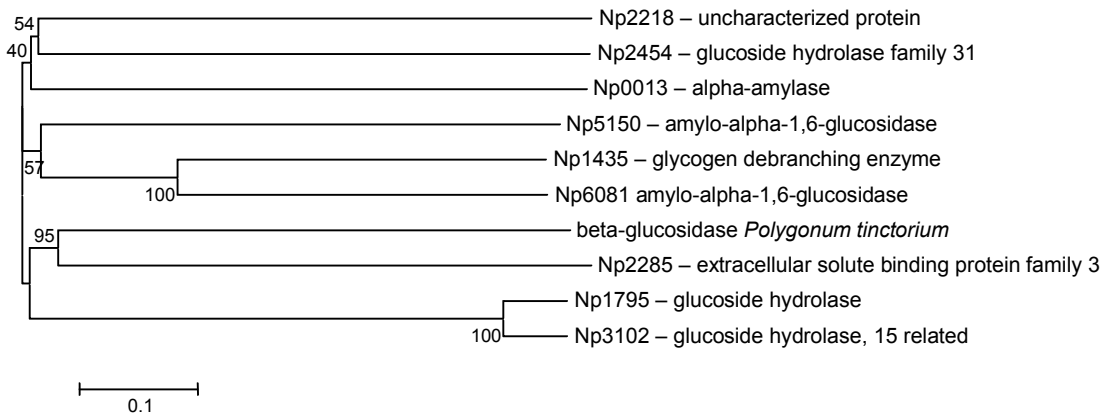


Figure 2.8: Unrooted minimum evolution phylogenetic tree of the genes related to sugar cleavage found in *N. punctiforme* ATCC 29133. The β-glucosidase from *Polygonum tinctorium* clades with the extracellular solute binding protein, Np2285.

Discussion

The unique combination of aromatic amino acid derived precursors suggested by the scytonemin chemical structure indicates that this metabolite is produced by an unusual mechanism unlike the common polyketide synthase (PKS) and non-ribosomal peptide synthase (NRPS) mechanisms often used by cyanobacteria (Dittmann *et al*., 2001). Due to the observed elicitation of scytonemin by UV radiation, preliminary theories on its biosynthesis considered whether a photochemical, enzymatic or a combination of mechanisms was responsible for its production. The inhibition of scytonemin production by chloramphenicol, a prokaryotic protein synthesis inhibitor, and its production after a lag period of two to five days, suggests that an enzymatic mechanism plays a role in the biosynthesis of this pigment (Garcia-Pichel and Castenholz, 1991). However, at the time of this study, the genes involved in the

biosynthesis of scytonemin were unknown. Insights into potential mechanisms involved in its biosynthesis were suggested through study of the scytonemin chemical skeleton. These insights resulted in a predicted mechanism for scytonemin biosynthesis as outlined in **figure 2.9**.

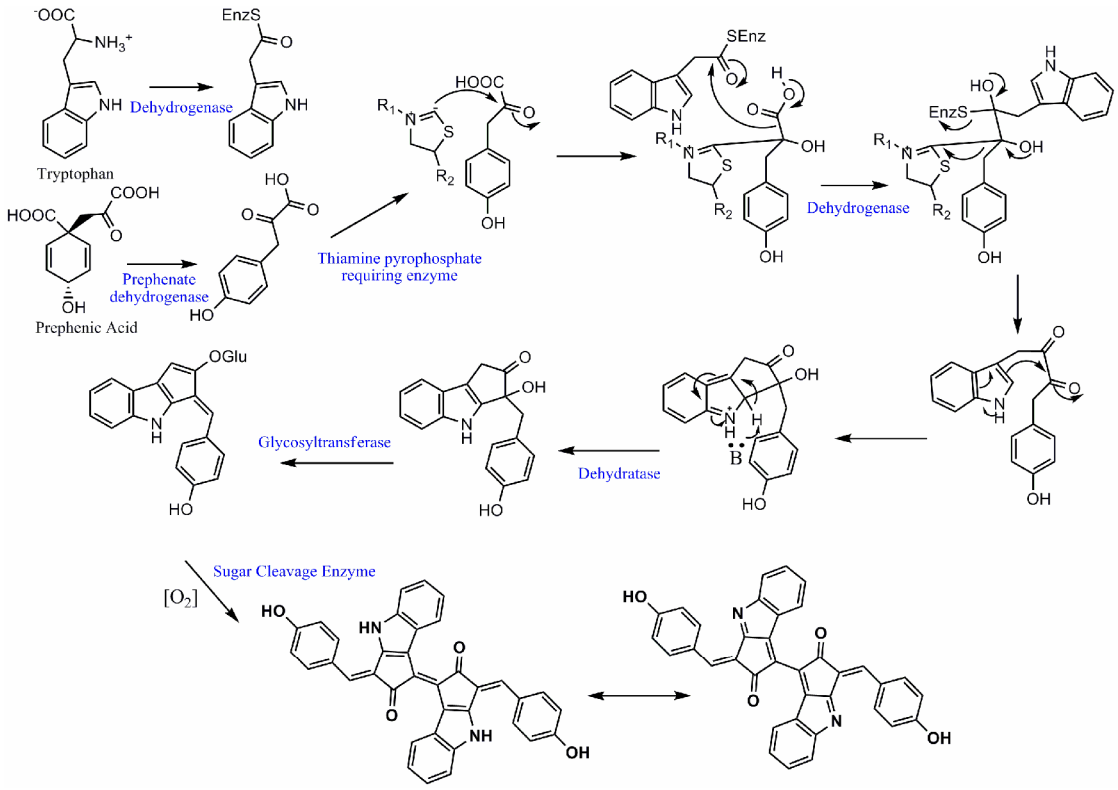


Figure 2.9: Diagram of predicted mechanism for scytonemin biosynthesis. Aromatic amino acid precursors result from aromatic amino acid biosynthesis as described in figure 2.1. Predicted enzymatic functions involved in biosynthesis are highlighted in blue.

Scytonemin is predicted to be derived from two aromatic amino acid precursors, a derivative of tyrosine such as 4-hydroxyphenylpyruvic acid and a derivative of tryptophan such as indole-3-acetic acid. Based on the use of these types of substrates, the genes involved in scytonemin biosynthesis are predicted to be related to the biosynthesis and metabolism of aromatic amino acids. A critical step in this biosynthesis is predicted to be the enzymatic facilitation of a decarboxylation reaction catalyzing the carbon-carbon bond formation between the two aromatic substrates. This type of decarboxylation may require the use of enzymatic cofactors such as pyridoxal-5-phosphate or thiamine pyrophosphate (Christen and Mehta, 2001; Jordan, 2003). In plants, aromatic amino acid decarboxylases are also involved in the biosynthesis of secondary metabolites, and are involved in interactions between the plant and its abiotic environment (Facchini *et al*., 2000). The prediction that an enzyme catalyzing the decarboxylation of the aromatic substrates is involved in the biosynthesis of scytonemin resulted in a close scrutiny of the *N. punctiforme* ATCC 29133 genome for candidate genes predicted to be involved in these types of reactions.

The formation of the diketone intermediate is predicted to be followed by an electron rearrangement and dehydration forming a monomeric scytonemin subunit. Dimeric scytonemin is found sequestered in the sheaths of cyanobacteria (Garcia-Pichel and Castenholz, 1991); therefore, the monomer is predicted to be transported into the sheath prior to dimerization. The transport of the monomer out of the cell may be facilitated by glycosylation of the monomer. Once the subunit is in the sheath,

the sugar moiety is cleaved and oxidative coupling would give rise to the dimeric form of scytonemin.

These latter steps in scytonemin biosynthesis are predicted to be similar to those seen during the biosynthesis of indigo. The biosynthesis of indigo begins with the enzymatic conversion of indole-3-glycerolphosphate to indole. Indole is then hydroxylated to form indoxyl. In the presence of oxygen, indoxyl will spontaneously form indigo (Warzecha *et al*., 2007). However, plants utilize an indoxyl-UDPG-glucosyltransferase to stabilize indoxyl by the addition of a sugar moiety to form indican (Marcinek *et al*., 2000; Warzecha *et al*., 2007). Indican is stored in the plant vacuole and is released after cell death. The glucosidases then cleave the sugar yielding indoxyl that can once again dimerize to form indigo (Warzecha *et al*., 2007). This mechanism is outlined in **figure 2.10**. *Polygonum tinctorium* is a plant cultivated to produce indigo. This plant was shown to have a specific β-glucosidase involved in hydrolyzing indican to yield indigo (Minami *et al*., 1999). Scytonemin biosynthesis may also require the use of a similar glucosidase to cleave the predicted glycoside involved in transport. Analysis of the *N. punctiforme* ATCC 29133 genome reveals that Np2285 aligns most closely to the β-glucosidase from *P. tinctorium* when compared to the other genes in the genome predicted to be involved in sugar cleavage (**Figure 2.8**). This β-glucosidase is not associated with any surrounding genes likely involved in aromatic amino acid biosynthesis and metabolism. However, a glucosidase involved in scytonemin biosynthesis would most likely be

compartmentalized to the sheath; therefore, its biosynthesis may not be directly associated with the scytonemin biosynthetic gene cluster.

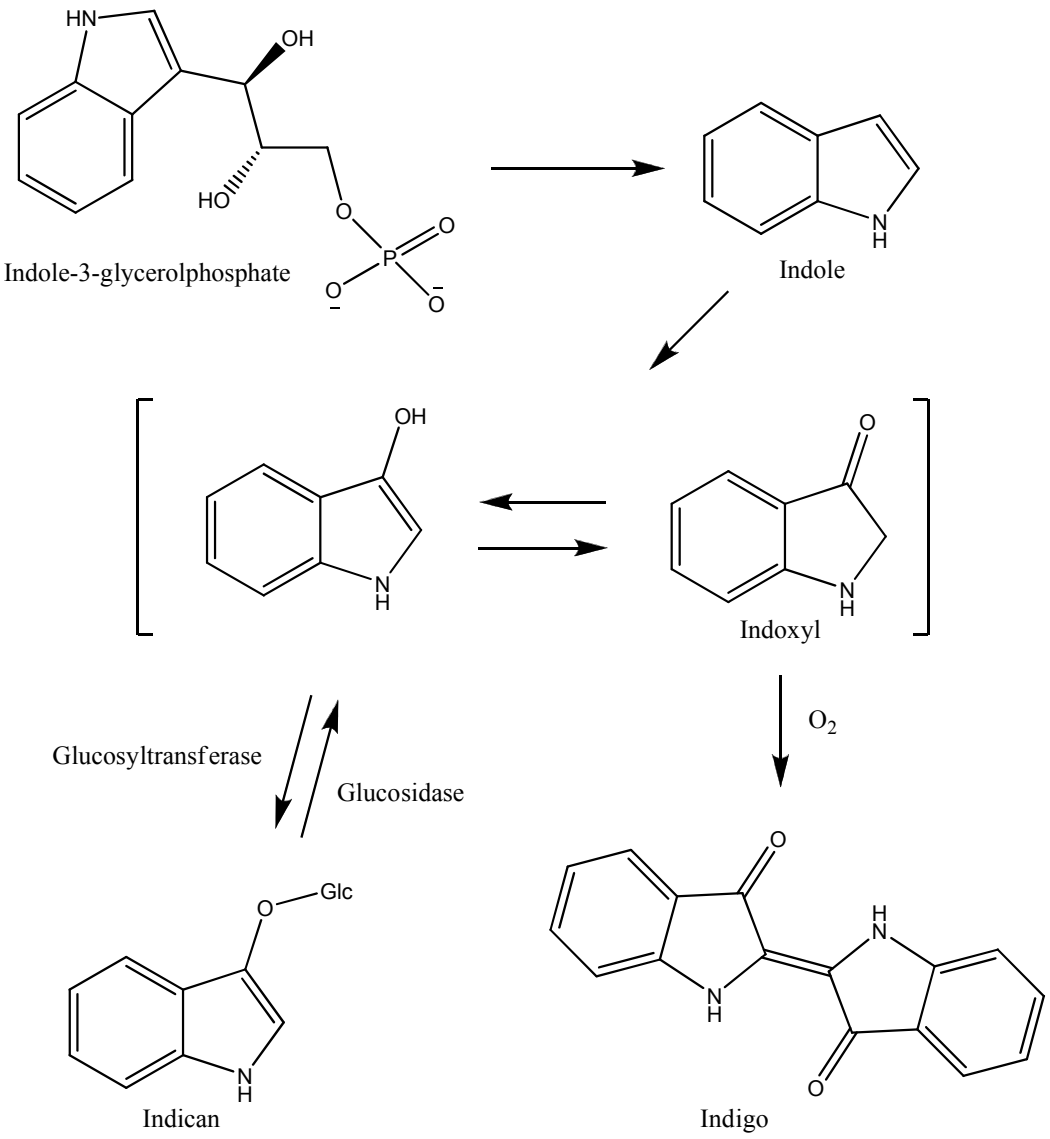


Figure 2.10: Diagram of indigo biosynthesis (Recreated from Warzecha *et al*., 2007).

The insights from this predicted mechanism revealed four candidate regions for involvement in scytonemin biosynthesis (**Figure 2.3**). Region NpR1802 contained genes predicted to be involved in signaling, transport, unknown functions, and hypothetical proteins; however, there was no evidence for involvement with aromatic amino acid biosynthesis or metabolism. Region NpR0978 contained a prephenate dehydrogenase as well as several genes with functions unrelated to the predicted biosynthesis. This region's involvement in scytonemin biosynthesis seemed unlikely due to the lack of genes related to tryptophan biosynthesis and metabolism. The third region, NpR6003, was a candidate due to the presence of a tyrosine/tryptophan transport protein. This region also contained genes predicted to be involved in signaling; however, there were no genes with functions related to aromatic amino acid metabolism.

The final region, NpR1263, contained many of the elements expected for scytonemin biosynthesis including aromatic amino acid biosynthesis, signaling, sugar transfer, unknown functions and hypothetical proteins. The most interesting characteristic of this region was the presence of NpR1263, a hypothetical protein with significant sequence similarity to a tyrosinase known to be involved in melanin biosynthesis in bacteria. Melanins are dark colored, insoluble pigments resistant to acid and susceptible to bleaching. These pigments are produced by many organisms for photoprotection (Krol and Liebler, 1998; Nosanchuk and Casadevall, 2003). Although the chemical structure of melanins in many organisms is unknown, it is believed to be a heterogenous polymer of mainly dihydroxyindole derived from

tyrosine and 3,4-dihydroxy-L-phenylalanine (DOPA) (Kotob *et al*., 1995; Krol and Liebler, 1998). In many cases, the formation of melanin utilizes enzymes involved in phenol coupling (Kotob *et al*., 1995; Williamson *et al*., 1998). Enzymes involved in phenol coupling include P450 monooxygenases, ascorbate oxidases, peroxidases, laccases and tyrosinases (Nezbedová *et al*., 2001). The identification of NpR1263 with sequence similarity to a tyrosinase involved in melanin biosynthesis and the role of both melanin and scytonemin in photoprotection presented an interesting correlation in predicted biosynthetic function. NpR1263 is also surrounded by genes predicted to be involved in aromatic amino acid biosynthesis, including all of the genes required for the biosynthesis of tryptophan. Therefore, the candidate region NpR1263 is the most promising region for involvement in scytonemin biosynthesis based on this bioinformatic analysis of the genome.

The elicitation of scytonemin by UV radiation suggests that the regulation of the genes involved in its biosynthesis would lead to a change in transcriptional level during exposure. Identification of candidate genes upregulated by UV radiation would provide further support for their involvement in scytonemin biosynthesis. Differential transcriptomics often involves the analysis of genes during two different environmental treatments (Frias-Lopez *et al*., 2004). In this study, we analyzed RNA from cultures without exposure to UV radiation and during exposure to UV radiation. This differential transcription profile can be analyzed through many methods including differential display, microarrays, suppressive subtractive hybridization (SSH), and semi-quantitative reverse transcription PCR (sqRT-PCR; Liang and

Pardee, 1997; Ji, *et al*., 2002; Frias-Lopez *et al*., 2004; Stowe-Evans *et al*., 2004). Due to lower numbers of false positives expected and lower costs, we attempted to utilize SSH and sqRT-PCR to support our bioinformatic analyses.

SSH (**Figure 2.6**) is a PCR based subtraction method that can selectively amplify differentially expressed cDNAs between two environmental samples. This method is valuable compared to other subtractive hybridization techniques because it does not require physical separation of the single stranded oligonucleotides from the double stranded oligonucleotides. The SSH method is based on the use of terminal inverted repeats as adaptors on cDNA that drives the selective suppression of common transcripts while normalizing the variations in the abundance of mRNA transcripts. This method has been shown to achieve up to 1,000-fold enrichment in differentially expressed cDNAs (Diatchenko *et al*., 1996). Although this method is commonly used in eukaryotes and other subtractive hybridization techniques are used to analyze prokaryotes, to my knowledge, no studies have used SSH to analyze changes in the transcriptome of prokaryotes after stimulation by environmental factors (Neilan and Pomati, 2004; Triplett *et al*., 2006). Unfortunately, the methodology developed for SSH in cyanobacteria during this study did not yield the results expected. The hybridization products were found to contain the sequences for the specific adaptors used; therefore, the subtraction may have still occurred. However, the inability to use primers to specifically amplify only mRNA resulted in hybridization products only derived from rRNA. mRNA represents less than 5% of the total RNA typically found in a bacterial cell, leading to the possibility that large quantities of rRNA may have

prevented the analysis of the much smaller quantities of mRNA (Griffiths, 2007). Further development of cDNA amplification of mRNA in cyanobacteria is necessary before this method will be effective for use in differential transcriptomics. However, the increasing availability of genomic information and the lowering cost of high throughput methods for differential expression analysis, such as microarrays, will make these other techniques more useful for similar analyses in the future.

Using the methodology developed for cDNA synthesis during SSH, a second differential expression technique was applied to provide insight into the candidate regions discovered through bioinformatics, namely sqRT-PCR. Based on a visual comparison of the resulting cDNA from sqRT-PCR, three of the gene regions predicted by bioinformatics were shown to have one gene upregulated by UV radiation after 24 and 48 hours of exposure. Further studies of the upregulation of the surrounding genes will provide evidence for the potential involvement of these gene regions in the biosynthesis of scytonemin (continued in chapter 3).

Conclusions

Bioinformatics is a powerful tool for the study of the biosynthesis of secondary metabolites in prokaryotic organisms including cyanobacteria (Udwary *et al*., 2007). An organism's genome represents the total genetic potential available, and an understanding of this potential can lead to new discoveries. These discoveries may include biosynthetic mechanisms important in developing drugs at lower costs and the discovery of novel metabolites with unique bioactivities. Although the genetic

potential of an organism is extremely important, the ability to understand how those genes respond to environmental stimulation is also valuable. This response can lead to a better understanding of "cryptic" biosynthetic pathways or insights into regulatory mechanisms that may have biotechnological value (Wenzel and Müller, 2005; Gross *et al*., 2007).

In this study, we identified four genetic regions with the potential to be involved in scytonemin biosynthesis. One of these regions, Np1263, was very promising based on its genetic architecture and its upregulation by UV radiation. In 2007, this genetic region was in fact shown through transposon mutagenesis to be responsible for scytonemin biosynthesis (Soule, *et al*., 2007). This finding supports the use of bioinformatics combined with genetic techniques such as differential expression to identify biosynthetic gene clusters involved in secondary metabolite production in cyanobacteria.

Table 2.2: Genes used for bioinformatic analysis of the *N. puncti*forme ATCC 29133 genome. Genes are separated by functional category and include the host organism, accession number and the number of resulting gene in the *N. punctiforme* ATCC 29133 genome. Continued.

Gene Function Category	Organism	Annotation	Accession Number	Number of Genes Identified
Signaling	*Synechocystis* sp. PCC6803	phytochrome/hisitidine kinase like	AAB39105	41
	Synechococcus sp. PCC7942	sensor kinase NblS	AAM18355	38
	Rhodopirellula baltica SH 1	light-regulated hisitidine kinase 1	NP_869840	22
	Nostoc sp. PCC7120	phytochrome-like protein	BAA88060	9
Tyrosine Metabolism	*Nostoc* sp. PCC7120	aspartate transaminase	BAB74039	5
	Bradyrhizobium sp. BTAi1	tyrosine transaminase	ZP_00859468.1	0
	Nostoc sp. PCC7120	histidinol-phosphate aminotransferase 1	Q8YV89	3
	Roseobacter sp. MED193	aromatic-amino-acid transaminase	ZP_01055262.1	0
	Nostoc sp. PCC7120	L-amino acid oxidase	NP_490275.1	3
	Rhodobacter capsulatus	tyrosine-phenol-lyase	AAC64208.1	1
	Ralstonia solanacearum GMI1000	Catechol oxidase	NP_518458.1	2
	Streptomyces avermitilis MA-4680	Monophenol monooxygenase	NP_828532.1	2
	Mesocricetus auratus	Tyrosine-3-monooxygenase	AAS79489.1	0
	Anabaena variabilis ATCC 29413	phenylalanine and histidine ammonia-lyase	YP_324488.1	1
	Sorghum bicolor	tyrosine N-monooxygenase	Q43135	3
	Lactococcus lactis	tyrosine Decarboxylase	CAF33980.1	0
	Silicibacter sp. TM1040	aromatic-L-amino acid decarboxylase	ZP_00623448.1	0

Table 2.2 continued…

Gene Function Category	Organism	Annotation	Accession Number	Number of Genes Identified
Tryptophan Biosynthesis	*Synechococcus* sp. WH8102	tryptophan synthase alpha subunit	NP_897612	2
	Synechococcus sp. WH8102	tryptophan synthase beta subunit	NP_898369	2
	Synechococcus sp. WH8102	indole-3-glycerol phosphate synthase	NP_897722	3
	Synechococcus sp. WH8102	putative anthranilate synthase component II	CAE07540	2
	Synechococcus sp. WH8102	anthranilate synthase component I/chorismate binding protein	NP_898136	3
	Synechococcus sp. WH8102	phosphoribosylanthranilate isomerase	NP_897824	1
Tryptophan Metabolism	*Trichodesmium erythraeum* IMS101	Tryptophan 2,3-dioxygenase	ZP_00675963	0
	Pseudomonas fluorescens Pf0-1	Tryptophan 2-monooxygenase	YP_350882	0
	Rubrobacter xylanophilus DSM9941	Aromatic-L-amino acid decarboxylase	ZP_00602286	0
	Nostoc sp. PCC 7120	L-amino acid oxidase	BAB78253	3
	Bacillus subtilis	Histidinol-phosphate/aromatic aminotransferase	P17731	4
	Prochlorococcus marinus str. MIT 9211	Tryptophan t-RNA ligase	ZP_01006688	1
	Xenopus laevis	Tryptophan 5-monooxygenase	Q92142	0
	Nostoc punctiforme PCC 73102	Tryptophanase	ZP_00109641	1
	Trichodesmium erythraeum IMS101	Tryptophan 2'-dioxygenase	ZP_00675963	0
	Neisseria meningitidis	Indolelactate dehydrogenase	AAB09666	6
	Desulfovibrio desulfuricans subsp. desulfuricans str. G20	Phenylpyruvate decarboxylase	YP_388823	3
	Salmonella enterica subsp. enterica serovar Typhi str. CT18	Indolepyruvate decarboxylase	NP_456948	3

Table 2.2 continued …

Gene Function Category	Organism	Annotation	Accession Number	Number of Genes Identified
Sugar Cleavage	*Polygonum tinctorium*	B-glucosidase (indigo)	BAA78708	1
	Synechococcus sp. WH 5701	B-glucosidase	ZP_01084621	1
	Synechococcus sp. WH 5701	Glycohydrolase	EAQ73943	0
	Prochlorococcus marinus str. MIT 9211	B-glucosidase	ZP_01006175	1
	Synechococcus elongatus PCC 7942	Glucosidase	YP_400591	1
Sugar Transfer	*Prochlorococcus marinus* str. MIT 9211	Sugar transferase	ZP_01004916	7
	Anabaena variabilis ATCC 29413	sugar transferase	YP_325324	7
	Synechococcus sp. WH 5701	glycosyltransferase	ZP_01086120	1
	Nostoc sp. PCC 7120	glycosyltransferase	NP_489280	2
	Gloeobacter violaceus PCC 7421	Glycosyltransferase	BAC91912	7
Oxidative Reaction	*Anabaena variabilis* ATCC 29413	Cytochrome P450	AAA80677	7
	Synechococcus elongatus PCC 6301	Cytochrome P450	YP_170898	7
	Trichodesmium erythraeum IMS101	Cytochrome P450	ZP_00674794	8
	Synechococcus sp. RS9917	Laccase	ZP_01081498	0
	Sinorhizobium meliloti	monophenol monooxygenase; tyrosinase	CAA49273	0
	Synechococcus elongatus PCC 7942	1-Cys peroxiredoxin	YP_401466	5
Transport	*Anabaena circinalis* 90	ABC transporter ATP-binding protein homologue	CAD60094	4

Acknowledgements

Chapter 2 is being prepared for publication of the material as a review. The dissertation author was the primary investigator and author of this material. I would like to acknowledge Dr. Aishu Ramaswamy for her assistance in learning bioinformatics and RNA isolation techniques in cyanobacteria.

References

Griffiths, A.J.F., Wessler, S.R., Lewontin, R.C., and Carroll, S.B. (2007) Proteins and their synthesis. In *Introduction to Genetic Analysis*. Macmillan. U.S.A. 319-350.

Christen, P. and Mehta, P.K. (2001) From cofactors to enzymes. the molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Chem. Rec*. 1: 436-447.

Diatchenko, L., Lau, Y.F.C., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., and Siebert, P.D. (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Nat. Acad. Sci.*. 93, 6025-6030.

Dillon, J.G. and Castenholz, R.W. (1999) Scytonemin, a cyanobacterial sheath pigment, protects against UVC radiation: implications for early photosynthetic life. *J. Phycol*. 35: 673-681.

Dillon, J.G., Tatsumi, C.M., Tandingan, P.G., and Castenholz, R.W. (2002) Effect of environmental factors on the synthesis of scytonemin, a UV-screening pigment, in a cyanobacterium (*Chroococcidiopsis* sp.). *Arch. Microbiol*. 177: 322-331.

Dittmann, E., Neilan, B.A., and Börner, T. (2001) Molecular biology of peptide and polyketide biosynthesis in cyanobacteria. *Appl. Microbiol. Biotechnol*. 57: 467-473.

Facchini, P.J., Huber-Allanach, K.L., and Tari, L.W. (2000) Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. *Phyotchem*. 54: 121-138.

Frias-Lopez, J., Bonheyo, G.T., and Fouke, B.W. (2004) Identification of differential gene expression in bacteria associated with coral black band disease by using RNA-arbitrarily primed PCR. *Appl. Environ. Microbiol*. 70: 3687-3694.

Garcia-Pichel, F. and Castenholz, R.W. (1991) Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycol*. 27: 395-409.

Garcia-Pichel, F., Sherry, N.D., and Castenholz, R.W. (1992) Evidence for an ultraviolet sunscreen role of the extracellular pigment scytonemin in the terrestrial cyanobacterium Chlorogloeopsis sp. *Photochem. Photobiol*. 56: 17-23.

Gross, H., Stockwell, V.O., Henkels, M.D., Nowak-Thompson, B., Loper, J.E., and Gerwick, W.H. (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chem. Biol*. 14: 53-63.

Hall, G.C., Flick, M.B., Gherna, R.L., and Jensen, R.A. (1982) Biochemical diversity for biosynthesis of aromatic amino acids among the cyanobacteria. *J. Bacterol*. 149: 65-78.

Herrmann, K.M. and Weaver, L.M. (1999) The shikimate pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol*. 50: 473-503.

Hunsucker, S.W., Klage, K., Slaughter, S.M., Potts, M., and Helm, R.F. (2004) A preliminary investigation of the *Nostoc punctiforme* proteome. *Biochem. Biophys. Res. Com*. 317: 1121-1127.

Ji, W., Wright, M.B., Cai, L., Flament, A., and Lindpaintner, K. (2002) Efficacy of SSH PCR in isolating differentially expressed genes. *BMC Genomics*. 3: 12.

Jordan, F. (2003) Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. *Nat. Prod. Rep*. 20:184-201.

Knaggs, A.R. (2003) The biosynthesis of shikimate metabolites. *Nat. Prod. Rep*. 20: 119-136.

Krol, E.S. and Liebler, D.C. (1998) Photoprotective actions of natural and synthetic melanins. *Chem. Res. Toxicol*. 11: 1434-1440.

Kotob, S.I., Coon, S.L., Quintero, E.J., and Weiner, R.M. (1995) Homogentisic acid is the primary precursor of melanin synthesis in Vibrio cholerae, a hyphomonas strain, and Shewanella colwelliana. *Appl. Environ. Microbiol.* 61: 1620-1622.

Krol, E.S. and Liebler, D.C. (1998) Photoprotective actions of natural and synthetic melanins. *Chem. Res. Toxicol*. 11: 1434-1440.

Liang, P. and Pardee, A.B. (1997) Differential display methods and protocols. In *Methods in Molecular Biology*. 85: 3-76.

Marcinek, H., Weyler, W., Deus-Neumann, B., and Zenk, M.H. (2000) Indoxyl-UDPG-glucosyltransfer from Baphicacanthus cusia. *Phytochem.* 53: 201-207.

Meeks, J.C., Elhai, J., Thiel, T., Potts, M., Larimer, F., Lamerdin, J., Predki., P., and Atlas, R. (2001) An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosyn. Res*. 70: 85-106.

Minami, Y., Shigeta, Y., Tokumoto, U., Tanaka, Y., Yonekura-Sakakibara, K., Oh-oka, H., and Matsubara, H. (1999) Cloning, sequencing, characterization, and expression of a β-glucosidase cDNA from the indigo plant. *Plant Sci*. 142: 219-226.

Neilan, B.A. and Pomati, F. (2004) PCR-based positive hybridization to detect genomic diversity associated with bacterial secondary metabolism. *Nucleic Acids Res.* 32, e7.

Nezbedová, L., Hesse, M., Drandarov, K., Bigler, L., Werner, C. (2001) Phenol oxidative coupling in the biogenesis of the macrocyclic spermine alkaloids aphelandrine and orantine in *Aphelandra* sp. *Planta*. 213: 411-417.

Nosanchuk, J.D. and Casadevail, A. (2003) The contribution of melanin to microbial pathogenesis. *Cell. Microbiol*. 5: 203-223.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y., and Hattori, M. (2001) Genome sequence of an industrial microorganism *Streptomyces avermitili*s: deducing the ability of producing secondary metabolites. *Proc. Nat. Acad. Sci*. 98: 12215-12220.

Proteau, P.J., Gerwick, W.H., Garcia-Pichel, F., and Castenholz, R. (1993) The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experentia*. 49: 825-829.

Sinha, R.P. and Häder, D.P. (2008) UV-protectants in cyanobacteria. *Plant Sci*. 174: 278-289.

Soule, T., Stout, V., Swingley, W.D., Meeks, J.C., and Garcia-Pichel, F. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J. Bacteriol*. 189: 4465-4472.

Sprenger, G.A. (2007) Aromatic amino acids. *Microbiol. Monogr*. 5: 93-127.

Stowe-Evans, E.L., Ford, J., and Kehoe, D.M. (2004) Genomic DNA microarray analysis: identification of new genes regulated by light color in the cyanobacterium *Fremyella diplosipon*. *J. Bacteriol*. 186: 4338-4349.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol*. 24:1596-1599.

Triplett, L.R., Zhao, Y., and Sundin, G.W. (2006) Genetic differences between blight-causing *Erwinia* species with different host specificities, identified by suppressive subtractive hybridization. *Appl. Environ. Microbiol.* 72, 7359-7364.

Warzecha, H., Frank, A., Peer, M., Gillam, E.M.J., Guengerich, F.P., and Unger, M. (2007) Formation of the indigo precursor indicant in genetically engineered tobacco plants and cell cultures. *Plant Biotech. J.* 5: 185-191.

Wenzel, S.C. and Müller, R. (2005) Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways. Curr. Opin. Biotechnol. 16: 594-606.

Williamson, P.R., Wakamatsu, K., and Ito, S. (1998) Melanin biosynthesis in Cryptococcus neoformans. *J. Bacteriol*. 180: 1570-1572.

Udwary, D.W., Zeigler, L., Asolkar, R.N., Singan, V., Lapidus, A., Fenical, W., Jensen, P.R., and Moore, B.S. (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Nat. Acad. Sci*. 104: 10376-10381.

CHAPTER THREE -

ORGANIZATION, EVOLUTION AND EXPRESSION ANALYSIS OF THE BIOSYNTHETIC GENE CLUSTER FOR SCYTONEMIN, A CYANOBACTERIAL ULTRAVIOLET ABSORBING PIGMENT

Abstract

Cyanobacteria are photosynthetic prokaryotes capable of protecting themselves from ultraviolet (UV) radiation through the biosynthesis of UV absorbing secondary metabolites such as the mycosporines and scytonemin. Scytonemin, a novel indolic-phenolic pigment, is found sequestered in the sheath where it provides protection to the subtending cells during exposure to UV radiation. The biosynthesis of scytonemin is encoded by a previously identified gene cluster that is present in six cyanobacterial species whose genomes are available. A comparison of these clusters reveals that two major cluster architectures exist which appear to have evolved through rearrangements of large sections such as those genes responsible for aromatic amino acid biosynthesis and through the insertion of genes that potentially confer additional biosynthetic capabilities. Differential transcriptional expression analysis demonstrated that the entire gene cluster is transcribed in higher abundance after exposure to UV radiation. This analysis helps delineate the cluster boundaries and indicates that regulation of this cluster is controlled by the presence or absence of UV radiation. The findings from an evolutionary phylogenetic analysis combined with the fact that the

scytonemin gene cluster is distributed across several cyanobacterial lineages leads to our proposal that the distribution of this gene cluster is best explained through an ancient evolutionary origin.

Introduction

Cyanobacteria are photosynthetic prokaryotes thought to be among the most ancient organisms on the planet (Cohen and Gurevitz, 2006). Their photosynthetic ability has long been speculated to have played a role in the oxygenation of the atmosphere, allowing for the development of many other life forms. However, before the presence of oxygen, cyanobacteria lived in an environment where the absence of a planetary ozone layer allowed exposure to high levels of harmful UV radiation (Wynn-Williams *et al*., 2002). The presence of high UV exposure levels early in the evolutionary history of cyanobacteria certainly presented these organisms with a major environmental pressure and resulted in the development of multiple UV defense adaptations which allow them to thrive in areas exposed to extremely high light and UV levels. These adaptations include avoidance, active repair mechanisms such as the SOS repair response, removal of reactive oxygen species by carotenoids, and biosynthesis of UV absorbing secondary metabolites, such as mycosporine amino acids and scytonemin. These adaptations are used in combination to avoid both acute cell damage (e.g., carotenoids) and the harmful effects of long-term UV radiation exposure (e.g., mycosporines and scytonemin; Ehling-Schulz and Scherer, 1999)

Scytonemin is an extracellular pigment first observed in 1849 when Nägeli

described a yellow-green pigmentation in the sheaths of cyanobacteria (Garcia-Pichel and Castenholz, 1991). In 1993, its chemical structure was elucidated and found to consist of an unprecedented dimeric indole-phenolic structure (Proteau *et al*., 1993). In pharmacological screens, this unique molecule was found to have both anti-inflammatory and anti-proliferative activity (Stevenson *et al*., 2002a; Stevenson *et al*., 2002b). Scytonemin is considered to be a true sunscreen agent due to its passive UV-absorption properties (Cockell and Knowland, 1999) in the UV-A region (*in vivo* λ_{max} = 370 nm; Proteau *et al*., 1993). Thus, 85-90% of the incident UV-A is absorbed by scytonemin in the sheaths of cyanobacteria, providing an effective protection to the subtending cells. Remarkably, this pigment has been described in over 300 species of cyanobacteria from various geographic locations and environments, and leads to intriguing questions concerning its evolutionary history (Garcia-Pichel and Castenholz, 1991).

In 2007, a cluster of genes involved in the biosynthesis of scytonemin was identified through the analysis of a non-scytonemin producing mutant of *Nostoc punctiforme* ATCC 29133 obtained through transposon mutagenesis (Soule *et al*., 2007). This mutation was embedded within a cluster of 18 open reading frames (ORFs NpR1259 to NpR1276) that were all transcribed in the same direction, thus suggesting this to be the functional genetic unit involved in scytonemin biosynthesis. This cluster contains genetic functions predicted to be involved in the biosynthesis of aromatic amino acids such as tryptophan, as well as other putative functions involved

in the assembly of scytonemin. However, information concerning the number of these genes that are transcribed during biosynthesis is lacking.

In this study, we provide evidence on the boundaries of the scytonemin biosynthetic gene cluster through a transcriptional expression analysis after exposure to UV radiation and through a comparison of the gene cluster as found in six cyanobacterial species. The conservation of this pathway across these cyanobacterial lineages also enabled an analysis of the evolution of these genetic elements, and provides support for the ancient origin of the scytonemin biosynthetic gene cluster.

Materials and Methods

Cyanobacterial Strains and Culture Techniques

The cyanobacterium *Nostoc punctiforme* ATCC 29133 was obtained from the American Type Culture Collection (ATCC). A culture was maintained in unialgal condition in liquid BG-11 freshwater media at 29^{o}C under a light intensity of approximately 19 μmol m^{-2} s^{-1} and a light/dark cycle of 16 h/8 h.

Transcriptional Expression Analyses

N. punctiforme ATCC 29133 was initially grown for approximately 45 days. Following this growth period, a portion of this culture was transferred to a petri dish and allowed to acclimated for 30 days prior to initiating the experiment. A sample was then taken from this dish culture as the –UV sample, and then the cyanobacterial filaments were exposed to UV radiation for 48 h (0.64 mW/cm^2; λ_{max} = 365nm). A second sample was taken from this dish after the 48 h exposure period, representing

the +UV sample. The extended acclimation time in the Petri dish cultures prior to taking samples for either the –UV or +UV samples controlled for pathway transcript levels between the two samples. Both –UV and +UV cultured biomass was extracted for RNA following a modified Trizol protocol (Invitrogen). The extracted RNA was treated with Turbo DNase (Ambion) and found to be free of genomic DNA contamination through control reverse transcription-PCRs (RT-PCRs) which lacked reverse transcriptase. RNA concentrations were determined using a Beckman Coulter DU800 spectrophotometer. Specific primers were designed to amplify regions of RNA ranging from 160 to 220 bp for 48 genes including those reported to be part of the scytonemin biosynthetic cluster and surrounding neighbors (Soule *et al*., 2007). Primer sequences are available at the end of the chapter in **table 3.3**.

Complementary DNA (cDNA) was synthesized following a standard protocol for Superscript III reverse transcriptase with an incubation at 50^{o}C (Invitrogen) and using 350 ng of RNA as template for every reaction. The PCR amplification was accomplished using a Taq polymerase mastermix (Promega), two microliters of first strand cDNA template and an annealing temperature of either 55 or 56^{o}C based on optimization for each primer set on genomic DNA. Genomic DNA was isolated and purified following a modified phenol protocol. For each of the 48 genes, four cDNA reactions were completed: -UV, +UV, negative control (no reverse transcriptase or RNA) and a positive control (created through PCR of genomic DNA). These four samples were visualized on a 1.5% agarose gel stained with ethidium bromide and documented using a Fisher Biotech transilluminator and a Kodak photographic system

with a 5 second exposure optimized for ethidium bromide. The visualization of each cDNA sample was repeated three times to minimize variations from gel loading by pipette. Analysis of gel band intensity was completed using the Biorad Quantity One software (Biorad).

For each cDNA band, an average band intensity was calculated and normalized for gel background by subtracting the average negative control intensity. In order to minimize the effects of primer bias, the DNA amplicons were normalized to one another by calculating a percentage of the positive control for each of the –UV and +UV bands. The percent band intensity of the -UV gel band was then subtracted from the percent band intensity of the +UV gel band and graphically represented using Excel software program (Microsoft). The standard error was calculated based on technical replicates by comparing the three gels run for each of the 48 primer sets, and any negative values are represented as zero values. An additional experimental replicate was also completed using the same methods described above and confirmed the results of the trend in increased expression discussed below.

Bioinformatic Analysis

Using the Joint Genome Institute (JGI) and National Center for Biotechnology Information (NCBI) web databases, 32 complete and 17 incomplete cyanobacterial genomes were examined for the scytonemin biosynthetic gene cluster using sequence similarity Blast searches (NCBI) with the ORFs Np1264 (S5) and Np1276 (S17) from the *N. punctiforme* ATCC 29133 scytonemin biosynthetic cluster (Soule *et al*., 2007). Five additional putative scytonemin biosynthetic gene clusters were located, and the

ORFs found in each of these were identified through available genome annotations as well as through manual identification using the NCBI ORF Finder. The six scytonemin gene clusters were each assembled using Vector NTI software program (Invitrogen), and the amino acid sequences for individual ORFs compared to one another using bl2seq Blast searches (NCBI) and through alignments created in Vector NTI.

Phylogenetic Analyses

Phylogenetic analyses were completed using MEGA 4.0 software program (Tamura *et al*., 2007). All sequence alignments were performed using ClustalW aligorithms with the Gonnet protein weight matrix for amino acids and the IUB DNA weight matrix for nucleotide based alignments. Sequence alignments were manually edited to exclude ambiguous regions. Neighbor-joining, minimum evolution, and maximum parsimony trees were created and evaluated with 10,000 bootstrap replicates. Outgroups used in the creation of the phylogenetic trees were obtained by identifying homologues by Blast searches. The basis of evolutionary selection for ORFs S6, S7, S16 and S17 was calculated using the number of nonsynonymous substitutions per nonsynonymous site (K_A) and the number of synonymous substitutions per synonymous site (K_S) in MEGA 4.0 (Tamura *et al*., 2007).

Results

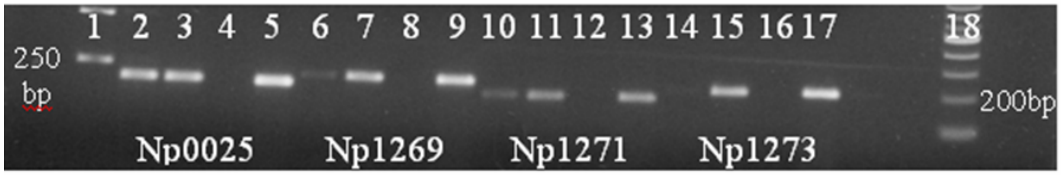
Transcriptional Expression of the Scytonemin Biosynthetic Pathway

The previously identified ORFs of the scytonemin biosynthetic gene cluster (Soule *et al*., 2007), as well as those ORFs from the surrounding region in the *N. punctiforme* ATCC 29133 genome, were analyzed for relative changes in their transcriptional levels after exposure to UV radiation. A total of 42 ORFs spanning 63.7 kb of the annotated genome from Np1239 to Np1288 were evaluated in this manner. Transcripts from the putative biosynthetic cluster (Np1260 to Np1276 and annotated in **Table 3.3** located at the end of the chapter) were substantially increased in cultures irradiated with UV radiation, as visualized by semiquantitative RT-PCR on agarose gels (**Figure 3.1**). Conversely, transcripts from outside of the reported cluster, Np1239 to Np1259, showed less than 10% increase in the UV stimulated cultures over the UV deficient cultures. This low level of intensity difference between +UV and –UV cDNA is similar to that for the gyrase control ORF Np0025. Because gyrase represents a housekeeping gene, it is not expected to show a transcriptional variation in response to UV radiation. The difference in the intensity of the +UV cDNA and that of the –UV cDNA for the ORFs Np1260 to Np1276 ranged from 44% for Np1265 to 82% for Np1274. ORF Np1277, a putative histidine kinase, showed an increase in intensity upon UV radiation stimulation by approximately 35%. ORFs Np1278 to Np1288 showed very modest increases in cDNA in the UV stimulated cultures, generally less than 20%. Thus, the major ORFs increased in expression by growth under UV radiation are those previously identified as the scytonemin biosynthetic

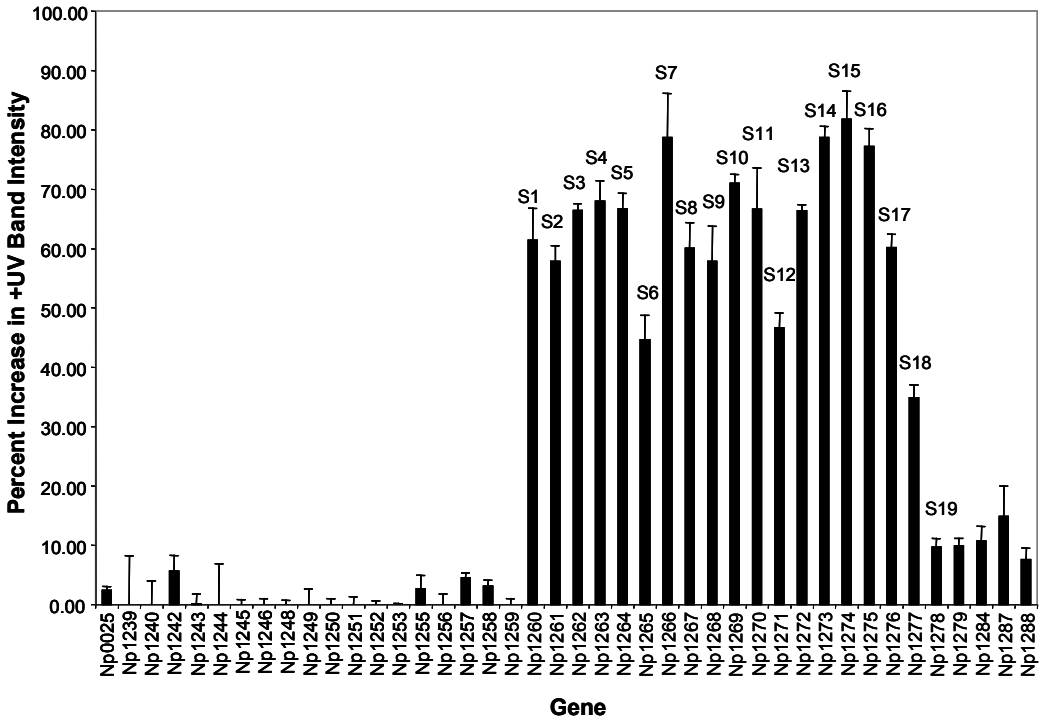
gene cluster with the addition of one downstream gene (Np1277) which is likely involved in regulation of the cluster.

Figure 3.1 - A. Example of *N. punctiforme* ATCC 29133 RT-PCR results for cDNA on a 1.5% agarose gel visualized with ethidium bromide. These data were used to create the graphical comparison in B. Lane 1 – 1kb ladder, lane 2 – Np0025 -UV cDNA, lane 3 – Np0025 +UV cDNA, lane 4 – Np0025 negative control, lane 5 – Np0025 genomic DNA positive control, lane 6 – Np1269 – UV cDNA, lane 7 – Np1269 +UV cDNA, lane 8 – Np1269 negative control, lane 9 – Np1269 genomic DNA positive control, lane 10 – Np1271 -UV cDNA, lane 11 – Np1271 +UV cDNA, lane 12 – Np1271 negative control, lane 13 – Np1271 genomic DNA positive control, lane 14 – Np1273 –UV cDNA, lane 15 – Np1273 +UV cDNA, lane 16 – Np1273 negative control, lane 17 – Np1273 genomic DNA positive control, lane 18 – 100bp ladder. B. Graphical comparison of the variation in transcriptional levels of the scytonemin biosynthetic pathway open reading frames and those open reading frames in the surrounding area in the *N. punctiforme* ATCC 29133 genome. This analysis shows the percent difference between the band intensity for the +UV cDNA over the –UV cDNA compared to the genomic DNA control. Each bar represents an individual gene in the pathway and Np0025 is the control representing gyrase from *N. punctiforme* L29133.

A


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
250 bp
200bp
Np0025
Np1269
Np1271
Np1273

B


Percent Increase in +UV Band Intensity
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
S13
S14
S15
S16
S17
S18
S19
Np0025 Np1239 Np1240 Np1242 Np1243 Np1244 Np1245 Np1246 Np1248 Np1249 Np1250 Np1251 Np1252 Np1253 Np1255 Np1256 Np1257 Np1258 Np1259 Np1260 Np1261 Np1262 Np1263 Np1264 Np1265 Np1266 Np1267 Np1268 Np1269 Np1270 Np1271 Np1272 Np1273 Np1274 Np1275 Np1276 Np1277 Np1278 Np1279 Np1284 Np1287 Np1288
Gene

Organization of the Scytonemin Biosynthetic Gene Cluster

Forty-nine cyanobacterial partial or complete genomes were analyzed for the scytonemin biosynthetic cluster using Np1276 as a query for Blast searches. Np1276 is a thiamine diphosphate-containing enzyme found within this cluster from *N. punctiforme* (Accession #NC_010628). As a result, five additional cyanobacterial species were found to contain a closely related gene cluster: *Nodularia spumigena* CCY9414 (Accession # NZ_AAVW01000004), *Cyanothece* sp. PCC7822 (Accession # NZ_ABVE01000001), *Cyanothece* sp. PCC7424 (Accession # ABOY01000036), *Lyngbya* sp. PCC8106 (Accession # AAVU01000005), and the previously identified gene cluster from *Nostoc* sp. PCC7120 (Accession # NC_003272; Soule *et al*., 2007). Many of the ORFs in the clusters from these other species are from the same conserved protein family and share a high percentage of amino acid sequence similarity (**Table 3.1**). However, the ORFs found both upstream and downstream of the scytonemin cluster from *N. punctiforme* have no significant sequence similarity to the regions surrounding the putative scytonemin cluster from these other species, thus further helping to delineate the boundaries of the cluster. The scytonemin biosynthetic gene clusters for all six cyanobacterial species are displayed in **Figure 3.2** where each ORF is annotated with respect to its involvement in similar biosynthetic functions across these six species (S1 to S34).

Table 3.1 - Best Blast results and amino acid comparison of the conserved open reading frames in the scytonemin biosynthetic gene cluster. Gene designations are assigned based on the original designation from the *N. punctiforme* ATCC 29133 gene cluster and are used uniformly throughout the five other identified gene clusters. The architecture of these genes across the six gene clusters is diagrammed in **figure 3.2**. The identity column represents the percent identity of the amino acid sequence to its nearest non-scytonemin gene cluster containing homolog by Blast analysis and is followed by the number of amino acids being compared in the ORF. The values in the column labeled 'Amino Acid Percent Identity for Conserved Proteins' were determined through the alignment of the ORFs from all six gene clusters.

Gene Designations	Gene Designations in *N. punctiforme*	Average Size (aa)	Conserved Protein Family	Best Blast Hit	Identity	Amino Acid Percent Identity for Conserved Proteins
S1	Np1260	355	AroA, 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase	carboxysome formation protein - *Cyanothece* sp. ATCC 51142	80% - 345	55%
S2	Np1261	368	TrpD, Anthranilate phosphoribosyltransferase	anthranilate phosphoribosyltransferase - *Anabaena variabilis* ATCC 29413	67% - 361	46%
S3	Np1262	416	TrpB, Tryptophan synthase beta chain	tryptophan synthase subunit beta - *Anabaena variabilis* ATCC 29413	90% - 394	70%
S4	Np1263	409	Tyrosinase	hypothetical protein DDBDRAFT_0185605 - *Dictyostelium discoideum* AX4	31% - 404	N/A
S5	Np1264	274	TrpA, Tryptophan synthase alpha chain	tryptophan synthase, alpha subunit - *Microcoleus chthonoplastes* pcc7420	73% - 263	54%

Table 3.1 continued …

Gene Designations	Gene Designations in *N. punctiforme*	Average Size (aa)	Conserved Protein Family	Best Blast Hit	Identity	Amino Acid Percent Identity for Conserved Proteins
S6	Np1265	287	TrpC, Indole-3-glycerol phosphate synthase	indole-3-glycerol-phosphate synthase - *Anabaena variabilis* ATCC 29413	70% - 262	44%
S7	Np1266	738	TrpE, Anthranilate/para-aminobenzoate synthases component I	anthranilate synthase - *Anabaena variabilis* ATCC 29413	77% - 726	55%
S8	Np1267	400	AroB, 3-dehydroquinate synthetase	3-dehydroquinate synthase - *Cyanothece* sp. CCY0110	66% - 389	45%
S9	Np1268	216	DsbA_FrnE, DsbA family, FrnE subfamily	hypothetical protein OA307_1438 - *Octadecabacter antarcticus* 307	36% - 203	70%
S10	Np1269	345	tyrA, bifunctional chorismate mutase/prephenate dehydrogenase	COG0287: Prephenate dehydrogenase - *Yersinia mollaretii* ATCC 43969	38% - 354	27%
S11	Np1270	430	Glycos_transf_1, Glycosyl transferases group 1	glycosyl transferase, group 1 - *Anabaena variabilis* ATCC 29413	74% - 426	77%
S12	Np1271	395	Uncharacterized conserved protein	NHL repeat-containing protein - *Candidatus Methanoregula boonei* 6A8	38% - 246	30%
S13	Np1272	453	None	conserved hypothetical protein - *Streptomyces ambofaciens*	28% - 423	21%
S14	Np1273	422	None	conserved hypothetical protein - *Streptomyces ambofaciens*	31% - 353	55%
S15	Np1274	325	None	phosphoenolpyruvate carboxykinase - *Heliobacterium modesticaldum* Ice1	29% - 120	52%
S16	Np1275	351	ELFV_dehydrog, Glutamate/Leucine/Phenylalanine/Valine dehydrogenase	leucine dehydrogenase - *Geobacillus thermodenitrificans* NG80-2	53% - 333	52%
S17	Np1276	625	TPP_enzymes, Thiamine diphosphate (TPP) enzyme family	acetolactate synthase large subunit - *Plesiocystis pacifica* SIR-1	51% - 586	41%
S18	Np1277	553	HATPase_c, Histidine kinase-like ATPases	PAS fold family - *Microcoleus chthonoplastes* PCC 7420	40% - 246	24%

Table 3.1 continued …

Gene Designations	Gene Designations in *N. punctiforme*	Average Size (aa)	Conserved Protein Family	Best Blast Hit	Identity	Amino Acid Percent Identity for Conserved Proteins
S19	Np1278	295	REC, Signal receiver domain	two component AraC family transcriptional regulator - *Anabaena variabilis* ATCC 29413	67% - 276	7%
S20	-	306	UbiA, 4-hydroxybenzoate polyprenyltransferase and related prenyltransferases	prenyltransferase, UbiA family - *Synechococcus* sp. PCC 7335	60% - 299	35%
S21	-	304	TatD_DNAse, TatD like proteins	hypothetical protein CY0110_32085 - *Cyanothece* sp. CCY0110	81% - 292	67%
S22	-	258	None	hypothetical protein S7335_3653 - *Synechococcus* sp. PCC 7335	52% - 234	39%
S23	-	460	Phosphodiest, Type I phosphodiesterase/nucleotide pyrophosphatase	hypothetical protein CY0110_09261 - *Cyanothece* sp. CCY0110	71% - 455	53%
S24	-	419	None	hypothetical protein CY0110_09256 - *Cyanothece* sp. CCY0110	59% - 396	40%
S25	-	423	None	PREDICTED: hypothetical protein - *Rattus norvegicus*	30% - 130	17%
S26	-	323	None	hypothetical protein MC7420_5057 - *Microcoleus chthonoplastes* PCC 7420	39% - 322	N/A
S27	-	451	Arginosuccinate synthase	conserved hypothetical protein - *Streptomyces ambofaciens* ATCC 23877	28% - 392	N/A
S28	-	130	None	Pep-cten putative exosortase interaction domain protein - *Microcoleus chthonoplastes* PCC 7420	54% - 35	N/A
S29	-	162	None	hypothetical protein RRC374 - uncultured methanogenic archaeon	36% - 87	N/A
S30	-	319	Qur: NADPH:quinone reductase and related Zn dependent oxidoreductases	zinc containing alcohol dehydrogenase superfamily protein - *Anabaena variabilis* ATCC 29413	50% - 318	N/A

Table 3.1 continued …

Gene Designations	Gene Designations in *N. punctiforme*	Average Size (aa)	Conserved Protein Family	Best Blast Hit	Identity	Amino Acid Percent Identity for Conserved Proteins
S31	-	117	None	hypothetical protein EHI_148590 - *Entamoeba histolytica* HM-1:IMSS	45% - 37	74%
S32	-	183	None	putative transposase - *Cyanothece* sp. PCC7425	59% - 44	N/A
S33	-	400	None	hypothetical protein RRC373 - uncultured methanogenic archaeon	21% - 208	N/A
S34	-	127	None	unnamed protein product - *Microcystis aeruginosa* PCC7806	73% - 123	83%

These six putative scytonemin clusters range in size from 33.2 kbp in *Cyanothece* sp. PCC7424 to 27.7 kbp in *Lyngbya* sp. PCC8106 (**Table 3.2**). There are 14 ORFs found conserved in all six cyanobacterial species, including the genes involved in aromatic amino acid biosynthesis (S1-S3, S5-S8 and S10), hypothetical proteins (S12-S13 and S15), a dehydrogenase (S16), a thiamine diphosphate-containing enzyme (S17), and a response regulator (S19). Five of these clusters contain an additional set of conserved ORFs (S20, S21, and S26) which encode for a prenyltransferase, a putative hydrolase, and a hypothetical protein.

Table 3.2 – Comparison of six cyanobacterial genomes found to contain the scytonemin biosynthetic gene cluster. *Strains of these species are reported to produce scytonemin.

Cyanobacteria	Genome				Scytonemin Pathway		
	Size, Mb	% G+C Content	Status	Sequence Location	Size, kb	Number of Open Reading Frames	Location in Genome
Nostoc punctiforme ATCC 29133*	8.2	41%	Complete	JGI	28	19	1516115 - 1544157
Nostoc sp. PCC7120 (=ATCC 27893)	6.4	41%	Complete	Kazusa DNA Research Institute	31.7	24	475640 - 507374
Nodularia spumigena CCY9414	5.3	42%	Incomplete	J. Craig Venter Institute	31.6	23	Unknown
Cyanothece sp. PCC7424	6.4	39%	Incomplete	JGI	33.2	25	Unknown
Cyanothece sp. PCC7822	5.7	40%	Incomplete	JGI	29.5	23	Unknown
Lyngbya sp. PCC 8106 (= *L. aestuarii* CCY9616)*	10	41%	Incomplete	J. Craig Venter Institute	27.7	21	Unknown

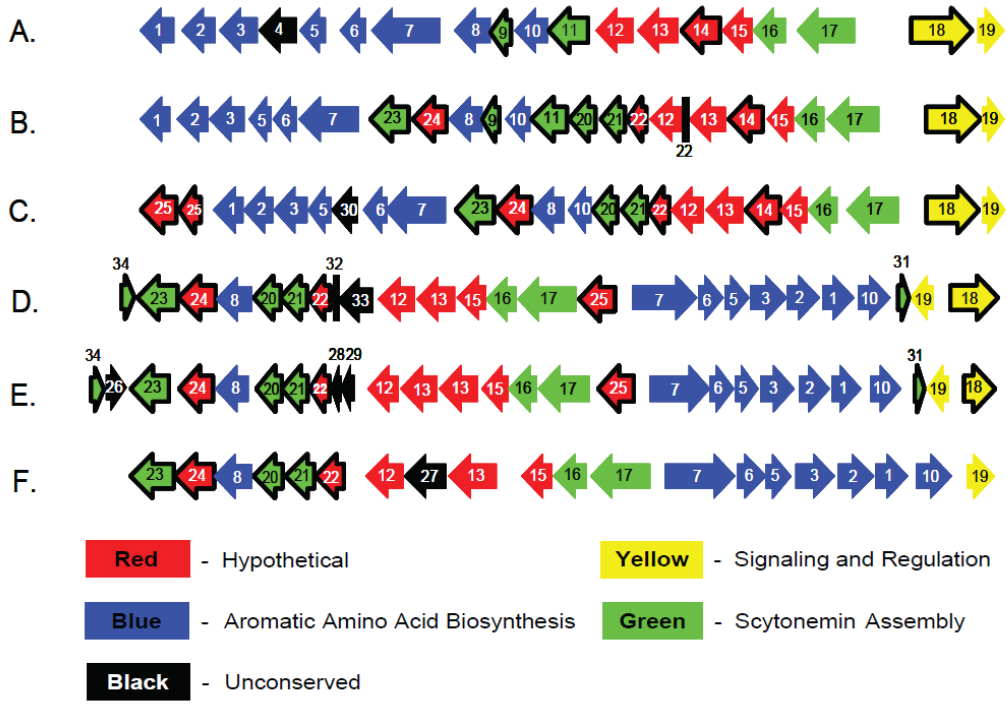


Figure 3.2 – Architectural comparison of open reading frames involved in the biosynthesis of scytonemin across six cyanobacterial species. A. *Nostoc punctiforme ATCC 29133*, B. *Nodularia spumigena CCY9414*,. C. *Nostoc sp. PCC7120*, D. *Cyanothece sp. PCC7822*, E. *Cyanothece sp. PCC7424*, F. *Lyngbya sp. PCC8106*. Arrows and lines represent individual ORFs with corresponding numbers annotated in table 3.1. Colors represent predicted gene functions as outlined in the key. Arrows with black outlines represent ORFs not conserved between the scytonemin pathways in all six of the cyanobacterial species.

Phylogenetic Analyses of the Biosynthesis of Scytonemin

The tryptophan synthase alpha subunit (TrpAα; 262aa), identified as Np1264 from the *N. punctiforme* ATCC 29133 genome, was used as a query for Blast searches against all 6 of the cyanobacterial genomes containing the scytonemin biosynthetic gene cluster. This analysis resulted in the identification of two TrpAα genes in each of these genomes. The amino acid sequences for all 12 of the TrpAα genes identified were used to construct a minimum evolution phylogenetic tree (**Figure 3.3**). The topology of this tree clearly resulted in the evolutionary separation of the TrpAα gene associated with each of the scytonemin biosynthetic gene clusters (blue) and the second TrpAα gene in each of the genomes (red).

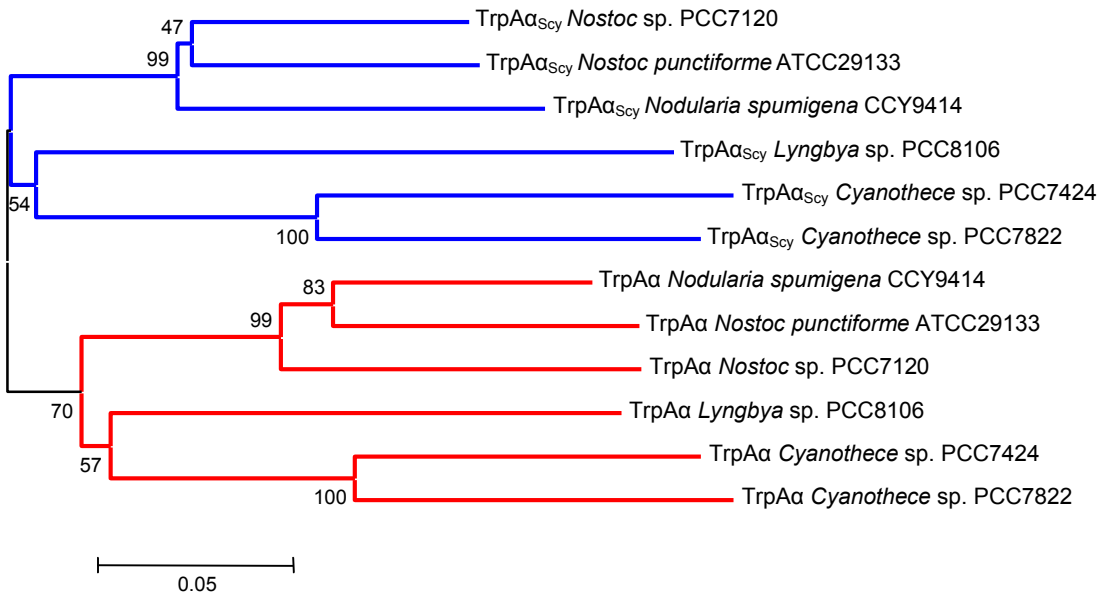


Figure 3.3: Phylogenetic comparison of the TrpAα genes from the six species of cyanobacteria containing the scytonemin biosynthetic gene cluster using minimum evolution criteria with bootstrap values greater than 45 labeled from 10,000 replicates. The blue lines represent the evolutionary separation of the TrpAα gene associated with the scytonemin biosynthetic gene cluster and the red lines represent the second TrpAα gene found in each of these genomes.

In a second phylogenetic analysis, two fused subsets were created *in silico* from the scytonemin gene cluster. The first contained 6,887 bp for the genes involved in aromatic amino biosynthesis (S1-S3, S5-S7 and S10) and the second contained 2,565 bp for the genes likely involved in scytonemin assembly (S16-S17). These were used to analyze the percent difference in mutations per site compared to the total number of amino acids found within each section of the scytonemin biosynthetic pathway (**Figure 3.4**). Based on a minimum evolution phylogenetic analysis, the genes encoding for aromatic amino acid biosynthesis showed approximately a two-fold increase in mutations per site compared to the scytonemin assembly genes.

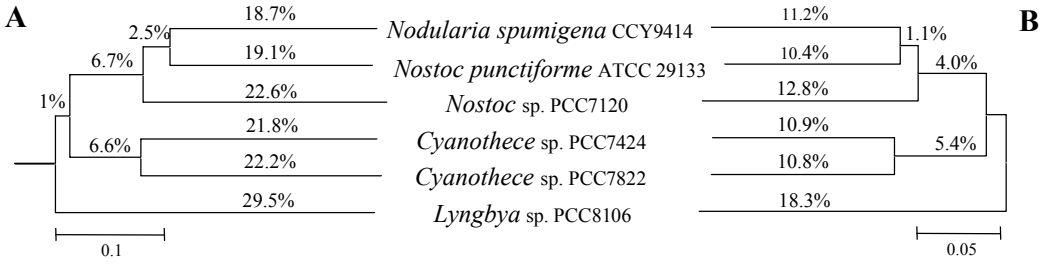


Figure 3.4 – Comparison of the sequence divergence between gene set S1-S3, S5-S7 and S10 involved in scytonemin-associated aromatic amino acid biosynthesis (A) and gene set S16-S17 conserved in the proposed scytonemin assembly portion of the cluster (B). Percentages are for the number of substitutions for each branch compared to the total number of amino acids used in the analysis. Phylogenetic trees were created using minimum evolution criteria and supported by neighbor-joining and maximum parsimony. Out groups used in the analysis are not shown (for S16-S17, *Anabaena variabilis* ATCC 29413 and for S1-S3,S5-S7 and S10, *Thermosynechococcus elongatus* BP-1).

In the final analysis, phylogenetic markers were created using 1,476 bp from the 16S rRNA and 1,859 bp from the rpoC1 gene, both of which are in common use for phylogenetic classification of cyanobacteria (Seo and Yokota, 2003), and compared to the sequence produced by fusing ORFs S6 - S7 (approximately 3129 bp) and S16 - S17 (approximately 3135 bp). These gene sets were used to evaluate phylogenetic congruence between the phylogenetic marker genes and those encoding scytonemin biosynthesis (**Figure 3.5**). Compared to the phylogenetic marker genes, all six scytonemin gene clusters appear to be fairly congruent with a similar branching pattern based on sequence divergence. The only differences appear to be the timing of the divergence and the variation in clustering of the two *Nostoc* species which is also apparent when the phylogenetic marker genes are compared.

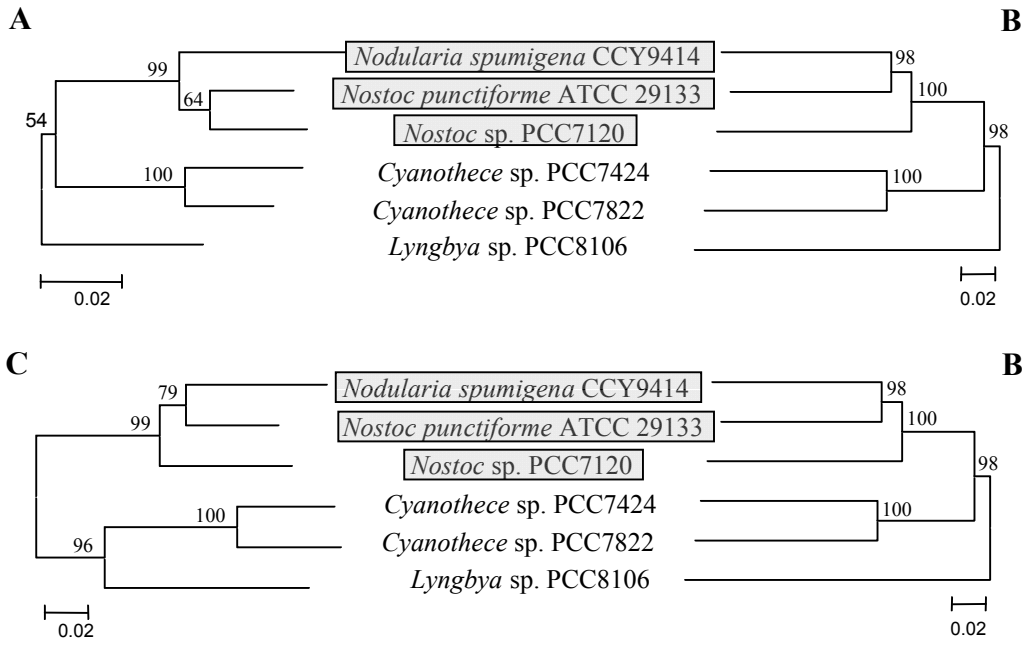


Figure 3.5 – Comparison of the evolution of phylogenetic markers, 16S rRNA (A) and rpoC1 (C), with four genes found in the scytonemin biosynthetic pathway (B), S6, S7, S16, S17. Phylogenetics are shown using minimum evolution criteria with bootstrap values from 10,000 replicates. These tree topologies are supported by both neighbor-joining and maximum parsimony criteria with bootstrap values greater than 50 for all branches (Results not shown). Grey boxes represent those cyanobacteria having the scytonemin gene cluster architecture where all ORFs are transcribed in one direction.

In order to determine if S6, S7, S16, and S17 are under positive or negative selection pressure, we compared the number of nonsynonymous substitutions per nonsynonymous site (K_A) to the number of synonymous substitutions per synonymous site (K_S) (Rantala *et al*., 2004). The ratios of K_A/K_S for these genes were below a value of 1.0 with an average of 0.473, indicating that this pathway is under purifying selection.

Discussion

The increase in the level of transcription for Np1260 through Np1276 in *N. punctiforme* ATCC 29133 after exposure to UV radiation is consistent with these genes forming the scytonemin biosynthetic gene cluster (**Figure 3.1**). Previous efforts utilized a mutagenesis approach of a single gene within this pathway to propose that these genes may form a biosynthetic functional unit (Soule *et al*., 2007); however, this report provides the first evidence that all 17 of these genes are expressed after exposure to UV radiation. This differential transcriptional analysis shows that there is an average 65% increase in the level of transcription for Np1260 through Np1276, which suggests that this gene cluster is under tight regulation and ultimately controlled by the presence of UV radiation. The regulation of the scytonemin gene cluster by an environmental parameter also suggests the presence of a unique signaling mechanism which responds to UV radiation. The presence of a conserved histidine kinase and response regulator just downstream of this cluster in five of the six identified pathways may indicate that this is the two-component signaling mechanism responsible for its

regulation. Although further studies, including quantitative PCR, are needed to better understand how variation in the wavelength of light affects this proposed regulatory histidine kinase, these semi-quantitative results indicate that the expression of this gene increases after exposure to ultraviolet radiation, although at a lower level than the other genes in the cluster. The apparent lower level of transcription may be a result of the phosphorylation kinetics of the two-component system. Because the phosphorylated state can have a half-life in the range of hours, this decreases the turnover time for proteins involved in the signaling response (West and Stock, 2001). In this study, the general increase in the level of transcription of the kinase and the presence of a conserved motif known as a Per Arnt Sim (PAS) fold, which has been shown to play a role in light induced regulation (Taylor and Zhulin, 1999), provides support for this hypothesis. A better understanding of UV induced signaling elements, such as the potential histidine kinase involved in scytonemin biosynthesis, is important to appreciating how organisms adapt to harsh light regimes. These signaling elements may also provide mechanistic insights into the transcriptional regulation of various classes of UV protectant molecules in other organisms and could have broad biotechnological applications as an inducible regulatory system.

Based on our differential transcriptional analysis, it appears that the enzymatic functions for the biosynthesis of scytonemin are found between Np1260 and Np1276. Previously, the biosynthesis of scytonemin was suggested to derive from a tryptophan-derived subunit and a phenylpropanoid-derived subunit (Proteau *et al*., 1993). The presence of eight genes involved in the biosynthesis of aromatic amino acids

(Np1260-Np1262, Np1264-Np1267 and Np1269) is consistent with this hypothesis. Interestingly, these eight aromatic amino acid biosynthetic genes are all clearly upregulated in the presence of UV radiation, indicating that their functionality is primarily for the biosynthesis of scytonemin. Previous work suggests that the scytonemin biosynthetic gene cluster could be considered a supraoperon due to the nesting of the tryptophan operon within this cluster (Xie *et al*., 2003b).

Consequently, the second copies of these aromatic amino acid biosynthetic genes, which are found scattered throughout these 6 cyanobacterial genomes, must be responsible for the primary metabolic needs of the cell. The use of these TrpAα homologs in separate metabolic functions is supported by the evolutionary relationship between these homologs across the 6 scytonemin gene cluster containing cyanobacterial lineages (**Figure 3.3**). The TrpAα genes that are associated with the scytonemin biosynthetic gene cluster associate with one another more closely than the homologs found within the same genome. The evolutionary separation of the two TrpAα genes from each of these genomes supports that the functional role of each of these gene is different, with one TrpAα gene functioning in primary metabolism (red lines) and a second functioning in specific secondary metabolism (e.g., scytonemin biosynthesis; blue lines). The scytonemin biosynthetic supraoperon thus encodes for a specialized metabolic capability that provides a selective advantage to cyanobacteria (Xie *et al*., 2003a).

Clarification of the boundaries for transcriptional regulation of the scytonemin biosynthetic gene cluster provides insights into some of the mechanistic steps that may

be involved in its biosynthesis in *N. punctiforme* ATCC 29133. For example, the conserved thiamine diphosphate-containing enzyme (Np1276) is likely involved in the decarboxylative coupling of two aromatic amino acid-derived precursors to form an acyloin intermediate. This reactivity has recently been confirmed through an *in vitro* study of the product of Np1276, which catalyzes the condensation of the two alpha-keto acids derived from tryptophan and tyrosine (Balskus and Walsh, 2008). Based on this *in vitro* data, the precursors are expected to couple in a similar manner *in vivo,* with the resulting acyloin intermediate subsequently cyclizing to the tetracyclic system present in each half of the scytonemin molecule (**Figure 3.6**). Dimerization to the completed scytonemin molecule may occur spontaneously or result from the tyrosinase activity (Np1263) found in the *N. punctiforme* ATCC 29133 gene cluster.

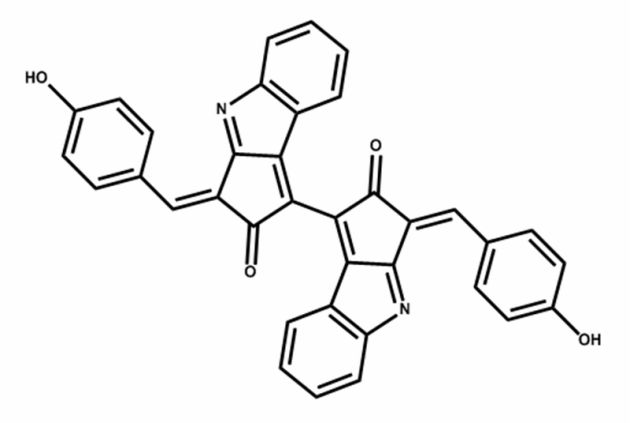


Figure 3.6 – Chemical structure of scytonemin

The putative scytonemin biosynthetic gene cluster is present in six species of cyanobacteria whose genomes are available (**Figure 3.2**). Based on the overall structure of these six gene clusters, they appear to have evolved by two major mechanisms: 1) the rearrangement of large sections of the cluster, and 2) the insertions of individual genes or small clusters of genes which confer additional biosynthetic capabilities. The two variations produced by rearrangements separate *Lyngbya* sp. PCC8106, *Cyanothece* sp. PCC7424, and *Cyanothece* sp. PCC7822 from the other three cyanobacterial species containing the scytonemin biosynthetic gene cluster. This rearrangement switches the transcriptional direction of S1-S3, S5-S7 and S10, which are involved in the biosynthesis of the aromatic amino acids, and moves them from upstream of S17 to downstream of this position (**Figure 3.7**).

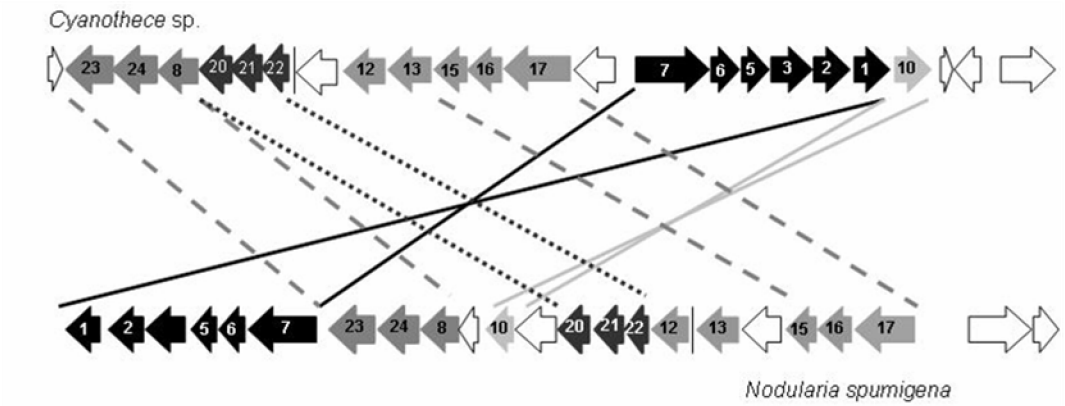


Figure 3.7 – Examples of the two major types of scytonemin biosynthetic gene clusters (top = *Cyanothece* sp. PCC7822 and bottom = *Nodularia spumigena* CCY9414). For purposes of clarity only, ORFs forming subclusters are shown in varying shades of white, grey and black. Lines between the clusters identify the rearrangement and inversion of ORFs involved in the biosynthesis of aromatic amino acids as well as other major differences between the two pathway architectures. Numbering of ORFs is based on annotations in **Table 3.1**.

This rearrangement is similar to what has been found in the microcystin biosynthetic gene cluster as identified in *Microcystis aeruginosa* and *Planktothrix agardhii*. The microcystin biosynthetic genes in *M. aeruginosa* are organized into two units that are transcribed in opposite directions whereas most of the microcystin genes in *P. agardhii* are transcribed in the same direction and as a single operon (Christiansen *et al*., 2003). The presence of these two types of rearranged scytonemin gene clusters suggests that this pathway must also involve two types of pathway promotion in order to regulate the production of this molecule. Based on the precedent formed by the microcystin biosynthetic gene cluster wherein there are two different promoters (Christiansen *et al*., 2003), one of these promoter regions may be a bidirectional promoter involved in initiating transcription of both the upstream and downstream parts of the scytonemin biosynthetic pathway.

The two architectures of the scytonemin biosynthetic gene cluster result in a splitting of the pathway into two subclusters, one containing the genes involved in the biosynthesis of aromatic amino acids and the other containing the putative scytonemin assembly genes. Based on a minimum evolution phylogenetic analysis of these two portions of the scytonemin gene cluster (**Figure 3.4**), the number of mutational changes between them is slightly different following the last evolutionary divergence. In this analysis, the genes involved in the biosynthesis of aromatic amino acids contain about 10% more substitutions per amino acid site than the scytonemin assembly genes. These results suggest that the S16 and S17 genes may be under a tighter selection pressure and less able to accommodate mutational changes than the genes involved in

aromatic amino acid biosynthesis. The weaker selection pressure imposed on the subcluster involving aromatic amino acid biosynthesis may be a result of the more dynamic properties reported for the trp operon (Xie *et al*., 2003b). It has been speculated that the trp operon has undergone a constant process of fine-tuning over evolutionary time because the operon encodes for biochemically expensive enzymatic reactions (Xie *et al*., 2003b).

Comparison of the six scytonemin gene clusters suggests that the insertion of individual genes or small clusters of genes has contributed to the evolution of the pathway. These inserted elements include an isomerase (S9) and glycosyltransferase (S11) in *N. punctiforme* ATCC 29133 and *N. spumigena* CCY9414 and a tyrosinase (S4) that is found only in *N. punctiforme* ATCC 29133. Because a set of genes (S20, S21, and S26) encoding a prenyltransferase, a putative hydrolase and a hypothetical protein are found in all of the scytonemin gene clusters except *N. punctiforme* ATCC 29133, we speculate that gene deletions are also involved in modifications to the pathway. However, there are a core set of genes involving aromatic amino acid biosynthesis (S1-S3, S5-S8 and S10) and scytonemin assembly (S12–S13 and S15-S17), which is highly conserved within all six gene clusters, and this likely signifies that these are key biosynthetic reactions for the production of UV sunscreen pigments. We speculate that the variable genes in these pathways represent enzymatic functions that are creating as yet uncharacterized structural variants of the scytonemin molecule with altered physiological or UV-screening properties.

A comparison of the 16S rRNA and rpoC1 data sets with that of four genes found in the scytonemin biosynthetic gene cluster, S6, S7 from the aromatic amino acid biosynthesis section and S16 and S17 from the scytonemin assembly portion, indicates that this pathway may be of ancient origin due to the congruence of the branching pattern between all of these species of cyanobacteria (**Figure 3.5**). These six gene clusters have a branching pattern similar to the phylogenetic marker genes except in the comparison of *N. punctiforme* ATCC 29133, *Nostoc* sp. PCC7120 and *N. spumigena* CCY9414, where variation is found when compared to the 16S rRNA evolutionary marker gene. However, this apparent evolutionary divergence may result from the fact that the 16S rRNA gene sometimes poorly defines fine-scale phylogenetic relationships of closely related species due to intragenomic variations between its multiple gene copies. In such cases, alternative phylogenetic marker genes, or use of gene combinations such as Multi-Locus Sequence Typing, can clarify these relationships (Case *et al*., 2007). These results are similar to the findings for the microcystin gene cluster where the biosynthetic sequences and the phylogenetic sequences (16S rRNA and rpoC1) have a high degree of congruence and similar levels of divergence; this is consistent with an ancient origin of the microcystins (Rantala *et al*., 2004). In the case of scytonemin, the congruence between the scytonemin data set and that of the evolutionary marker genes are also indicative of the ancient origin of this gene cluster.

The evolution of the scytonemin biosynthetic gene cluster from an ancient origin is further supported by its being under purifying selection pressure. A high

K_A/K_S ratio would indicate positive selection pressure due to the helpful influence of these mutations on an organism's fitness; however, the observed low K_A/K_S ratio (0.473) indicates a purifying selection pressure, which signifies that mutations affecting the protein sequence are deleterious and negatively affect an organism's fitness (Rantala *et al*., 2004). This purifying selection pressure is consistent with an ancient origin of the scytonemin biosynthetic gene cluster (Massingham and Goldman, 2005).

The tight transcriptional regulation of the scytonemin biosynthetic gene cluster by UV radiation and its apparent ancient evolution are evidence to this molecule's ecological significance as a sunscreen pigment. The idea that cyanobacteria would support a biosynthetic pathway of high energetic requirements, due to the use of aromatic amino acids, throughout their evolution even after the formation of the protective ozone layer is further support for the importance of scytonemin in expanding the habitat of these diverse organisms. The presence of the scytonemin biosynthetic gene cluster across cyanobacterial lineages and the variations between individual pathways indicate that scytonemin is an adaptive and functional sunscreen, which allows the expansion of cyanobacteria into many different ecological niches exposed to high levels of UV radiation. Indeed, in the current era wherein greater levels of UV irradiation are reaching the earth's surface, this adaptation is of considerable significance to cyanobacterial survival and growth in new habitats. Finally, a better understanding of the signaling and biosynthetic mechanisms involved in this widespread and naturally-occurring sunscreen molecule may allow the

development of important biotechnological advances in both drug discovery and agricultural biology.

Table 3.3: Blast analysis of genes involved in scytonemin biosynthesis and those genes surrounding the gene cluster from the *N. punctiforme* ATCC 29133 genome. Columns include best Blast hit, sequence identity to the best hit, and primers and annealing temperatures used for differential transcription analysis.

Gene	Strand Direction	Size (aa)	Start Site	Best Blast Hit	Identity	Forward Primer Sequence (5'-3')	Reverse Primer Sequence (5'-3')	Optimal Annealing Temperature (°C)
Np0025	F	871	39862	DNA gyrase subunit A - *Anabaena variabilis* ATCC 29413	88	CCATGAG CGTGATA GTGG	CCAAGGCA TCATATAC TGC	55
Np1239	F	129	1492767	Hypothetical protein Ava_0805 - *Anabaena variabilis* ATCC 29413	87	CAGCGTC ATATCGA TCTCG	GTTCTTTC GCCGTCCT CG	56
Np1240	R	245	1493232	Protein of unknown function DUF477 - *Anabaena variabilis* ATCC 29413	75	CGAACAG TTACAATT CTGAC	CCTCGCTG TTTGCCGC A	56
Np1242	R	611	1494578	Peptidase S49, protease IV - *Anabaena variabilis* ATCC 29413	75	CCAACAA TTCGCGG ATCG	GCTCGTCT AAAGATAC AGG	55
Np1243	R	349	1496487	Isopentenyl pyrophosphate isomerase - *Anabaena variabilis* ATCC 29413	81	GCGACTT GGGCCAA ACG	GTTCAGTC TCATCAAA TCAC	55
Np1244	F	195	1497690	Protein of unknown function DUF820 - *Anabaena variabilis* ATCC 29413	63	GCTACTA CACCCCA GAAG	CACATCGC CTATGTAA ATCT	55
Np1245	R	126	1498429	Hypothetical protein Saro_3152 - *Novosphingobium aromaticivorans* DSM 12444	60	GACCAGA CGGTGAA CCAG	CGCCCTTG TTGTTGTT GGA	55
Np1246	R	83	1499009	Hypothetical protein Ava_2256 - *Anabaena variabilis* ATCC 29413	41	GAGTCTG CGGTGGA TTTTAACC	TGCCATTG ACTCAGCA AAG	55
Np1248	R	94	1499959	None		GAGCGAT GTCTACG AGGG	GTTTCAGG ACTGGAAT ACAC	55
Np1249	F	202	1500396	Ribosomal protein S8 - *Bacillus selenitireducens* MLS10	29	GAATGTA CTCAGTCT CGTAG	GTTTTAAA GTTTTGCC AGCATC	55

Table 3.3 continued …

Gene	Strand Direction	Size (aa)	Start Site	Best Blast Hit	Identity	Forward Primer Sequence (5'-3')	Reverse Primer Sequence (5'-3')	Optimal Annealing Temperature (°C)
Np1250	F	1079	1500944	HEAT repeat-containing PBS lyase - *Anabaena variabilis* ATCC 29413	44	TGAAGAG AAGCAGC GATTTG	TGCCAAAG AAATCCAA ATTG	56
Np1251	R	504	1504451	Hypothetical protein Ava_0218 - *Anabaena variabilis* ATCC 29413	82	TCACTACC CTCTACCC CTGC	CAGTTTCC TACTTTGG CGC	56
Np1252	F	564	1506176	Serine/Threonine protein kinase - *Anabaena variabilis* ATCC 29413	75	AGAACAA CGCGAGA ACGAAG	GCAATTGC TTGACACC ACC	55
Np1253	F	715	1507928	Serine/Threonine protein kinase - *Anabaena variabilis* ATCC 29413	73	TTAGCCG AAGATGT CAACCG	TTAGCCGC AGCAGTTC TCG	56
Np1255	F	123	1511789	YrnH - *Rhizobium* sp. NGR234	35	CTTCATTT TTCACAAT TACC	CTTTGGCG AATAAAAG TAA	55
Np1256	R	324	1512214	Pyridoxal-5'-phosphate-dependent enzyme, beta subunit - *Anabaena variabilis* ATCC 29413	85	GAAAAGA GCGGGTA GCATCG	TCATCCCC ATGTAGTC GCC	55
Np1257	F	270	1513397	Short-chain dehydrogenase/reductase SDR - *Anabaena variabilis* ATCC 29413	68	GTGCTGG TTAACAA TGCTGG	AAAAGCTG CTGCTGAG GCC	55
Np1258	F	66	1514456	Heavy metal transport/detoxification protein - *Anabaena variabilis* ATCC 29413	60	ATGACAA TTGAATTG AAAGTTC	GCCTTCTA TCGTATAA CCAG	55
Np1259	R	314	1514904	Signal transduction histidine kinase (STHK), LytS - *Anabaena variabilis* ATCC 29413	53	ATTCTTCA ATACCCC AACGC	GCTACTGG CACTTTTG CTGG	55
Np1260	R	366	1516115	Phospho-2-dehydro-3-deoxyheptonate aldolase, subtype 2 - *Crocosphaera watsonii* WH 8501	79	AGGAATC TGTACCA GCTGCG	GCGGGGAA TTAGAACC TTTG	55
Np1261	R	363	1517476	Anthranilate phosphoribosyltransferase - *Anabaena variabilis* ATCC 29413	67	GTGGAAG CGCCATCT CC	GTTGCTGA ATCGGCAA TCG	55
Np1262	R	409	1518683	Tryptophan synthase subunit beta - *Anabaena variabilis* ATCC 29413	89	GTTTAGC GAGTAAC ACCTG	CGCAACGA GCCGAGTT TC	55
Np1263	R	408	1519970	Hypothetical protein DDBDRAFT_0185605 - *Dictyostelium discoideum* AX4	32	TTGTGTCC GTAAGGC TGACC	TAGACGCA AGTACGAC ACCC	55

Table 3.3 continued …

Gene	Strand Direction	Size (aa)	Start Site	Best Blast Hit	Identity	Forward Primer Sequence (5'-3')	Reverse Primer Sequence (5'-3')	Optimal Annealing Temperature (°C)
Np1264	R	275	1521314	Tryptophan synthase subunit alpha - *Anabaena variabilis* ATCC 29413	73	CACTACC AACCCTT GCAC	GCAGATGG GCCTGTTA TTC	55
Np1265	R	275	1522619	Indole-3-glycerol-phosphate synthase - *Anabaena variabilis* ATCC 29413	70	GCACAGT AGGGGTA ATGC	CCTAGCTT AATTGCTG AGG	56
Np1266	R	734	1523627	Glutamine amidotransferase of anthranilate synthase - *Anabaena variabilis* ATCC 29413	77	CAAATCG CGCTGAT CTTTAG	GTCGGCTT TGTTAAAT CTAC	55
Np1267	R	394	1526310	3-dehydroquinate synthase - *Deinococcus geothermalis* DSM 11300	57	GCAGCAG CATATCCT ACC	GAAGCAAT TAGTTGCG TATAC	55
Np1268	R	214	1527491	Similar to frnE protein - *Mesorhizobium loti* MAFF303099	35	CAACTGC ATCAACG ACATCG	TTGCATTA GCACCGAC AAAC	56
Np1269	R	359	1528271	COG0287: Prephenate dehydrogenase - *Yersinia mollaretii* ATCC 43969	38	CGTATGTT CAATAGG AACAC	GTTGTCAT GCCAGCCT TAC	55
Np1270	R	429	1529365	Glycosyl transferase, group 1 - *Anabaena variabilis* ATCC 29413	74	AGCCAAA GCACCTG CTGAC	GCAGATGT GGGGTTGA TTG	56
Np1271	R	398	1530908	NHL repeat containing protein - *Candidatus methanoregula boonei* 6A8	32	TTTATAGC TTGTGCCA CCCG	AGCCGCGT ACAGGTGT TTG	55
Np1272	R	432	1532261	Hypothetical protein Adeh_0026 - *Anaeromyxobacter dehalogenans* 2CP-C	26	ATACCGC CCTGTTGA ATTTC	TATCCAGA AGGTGAAG CGC	56
Np1273	R	423	1533657	Conserved hypothetical protein - *Streptomyces ambofaciens* ATCC 23877	30	ATAGTCA AGGCGCT AGGCG	CCAACAAT CTTCGCCG ATG	55
Np1274	R	322	1535003	None		GCGAAAG CATAAGC GGTTG	CTGTTACG CCCAAATG GTG	56
Np1275	R	353	1535997	Leucine dehydrogenase - *Bacillus halodurans* C-125	52	GCTGCAC CAGCAAT AATCTG	ATCCAGCA AAAGCGGA AG	56

Table 3.3 continued …

Gene	Strand Direction	Size (aa)	Start Site	Best Blast Hit	Identity	Forward Primer Sequence (5'-3')	Reverse Primer Sequence (5'-3')	Optimal Annealing Temperature (°C)
Np1276	R	624	1537418	Acetolactate synthase large subunit - *Plesiocystis pacifica* SIR-1	51	CTAACAA CCACGAG CGCTTC	TCAGCACA GCCGTGAA ATAC	55
Np1277	F	659	1541129	Multi-sensor signal transduction histidine kinase - *Anabaena variabilis* ATCC 29413	38	ACTTTCCA TGAGGCT GCAAG	CGAAAGCA CAGCCAAA TTC	56
Np1278	F	283	1543306	Two component transcriptional regulator, AraC family - *Anabaena variabilis* ATCC 29413	67	CCGACTT AGTGATTT GCG	GATAGCGA TCGCTCTG AG	55
Np1279	R	400	1544295	Hypothetical protein Ava_2750 - *Anabaena variabilis* ATCC 29413	81	GCGATGA TCTGGTG ATTGA	CTGTTGTA TAGTGACT CGC	56
Np1284	R	551	1549256	ABC transporter ATP-binding protein - *Thermosynechococcus elongatus* BP-1	50	GTTCTAA GTGACGT AAGGG	GCTGCTGG TCTAATTG CC	55
Np1287	R	625	1553189	Conserved hypothetical protein - *Chlorobium phaeobacteroides* BS1	27	GTAGCTA TTAGGCA AAGGTT	CAGATATT GTGGCATT TCTG	55
Np1288	R	444	1555132	JamQ - *Lyngbya majuscula*	49	CGCAGTA AACCTTTG CTAC	CTCTGACA CAGGCTTT AGA	55

Acknowledgements

Chapter 3, in part, has been submitted for publication of the material as it will appear in Applied and Environmental Microbiology. 2009. Sorrels, C.M., Proteau, P.J., and Gerwick, W.H.. The dissertation author was the primary investigator and author of this paper.

This work was supported by a National Institutes of Health predoctoral fellowship through the Training Grant in Marine Biotechnology T32GM067550 and previous support from the Oregon Sea Grant program R/BT-40. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

We would like to acknowledge E. Esquenazi and P. Dorrestein at the University of California, San Diego for help identifying scytonemin production by MALDI MS, and J. Wells and family for general support to the laboratory. We also acknowledge L. Gerwick for helpful discussions on phylogenetics and molecular biological techniques.

References

Balskus, E. P., and C. T. Walsh. (2008) Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J. Am. Chem. Soc*. 130:15260-15261.

Case, R. J., Y. Boucher, I. Dahllöf, C. Holmström, W. F. Doolittle, and S. Kjelleberg. (2007) Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol*. 73:278-288.

Christiansen, G., J. Fastner, M. Erhard, T. Börner, and E. Dittmann. (2003) Microcystin biosynthesis in *planktothrix*: genes, evolution, and manipulation. *J. Bacteriol*. 185:564-572.

Cockell, C. S., and J. Knowland. (1999) Ultraviolet radiation screening compounds. *Biol. Rev. Camb. Philos. Soc*. 74:311-345.

Cohen, Y., and M. Gurevitz. (2006) The cyanobacteria – ecology, physiology and molecular genetics. *Prokaryotes.* 4:1074-1098.

Ehling-Schulz, M., and S. Scherer. (1999) UV protection in cyanobacteria. *Eur. J. Phycol*. 34:329-338.

Garcia-Pichel, F., and R. W. Castenholz. (1991) Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycol.* 27:395-409.

Massingham, T., and N. Goldman. (2005) Detecting amino acid sites under positive selection and purifying selection. *Genetics*. 169:1753-1762.

Proteau, P. J., W. H. Gerwick, F. Garcia-Pichel, and R. Castenholz. (1993) The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia*. 49:825-829.

Rantala, A., D. P. Fewer, M. Hisbergues, L. Rouhiainen, J. Vaitomaa, T. Börner, and K. Sivonen. (2004) Phylogenetic evidence for the early evolution of microcystin synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 101:568-573.

Seo, P. S., and A. Yokota. (2003) The phylogenetic relationships of cyanobacteria inferred from 16S rRNA, gyrB, rpoC1 and rpoD1 gene sequences. *J. Gen. Appl. Microbiol*. 49:191-203.

Soule, T., V. Stout, W. D. Swingley, J. C. Meeks, and F. Garcia-Pichel. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J. Bacteriol*. 189:4465-4472.

Stevenson, C. S., E. A. Capper, A. K. Roshak, B. Marquez, C. Eichman, J. R. Jackson, M. Mattern, W. H. Gerwick, R. S. Jacobs, and L. A. Marshall. (2002) The identification and characterization of the marine natural product scytonemin as a novel antiproliferative pharmacophore. *J. Pharmacol. Exp. Ther*. 303:858-866.

Stevenson, C. S., E. A. Capper, A. K. Roshak, B. Marquez, K. Grace, W. H. Gerwick, R. S. Jacobs, and L. A. Marshall. (2002) Scytonemin – a marine natural product inhibitor of kinases key in hyperproliferative inflammatory diseases. *Inflamm. Res*. 51:112-114.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol*. 24:1596-1599.

Taylor, B. L., and I. B. Zhulin. (1999) PAS domains: internal sensors of oxygen, redox potential and light. *Microbiol. Mol. Biol. Rev.* 63:479-506.

West, A. H., and A. M. Stock. (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *TRENDS Biochem. Sci.* 26:369-376.

Wynn-Williams, D. D., H. G. M. Edwards, E. M. Newton, and J. M. Holder. (2002) Pigmentation as a survival strategy for ancient and modern photosynthetic microbes under high ultraviolet stress on planetary surfaces. *Int. J. Astrobiol*. 1:39-49.

Xie, G., C. A. Bonner, T. Brettin, R. Gottardo, N. O. Keyhani, and R. A. Jensen. (2003a) Lateral gene transfer and ancient paralogy of operons containing redundant copies of tryptophan-pathway genes in *Xylella* species and in heterocystous cyanobacteria. *Genome Biol*. 4:R14.

Xie, G., N. O. Keyhani, C. A. Bonner, and R. A. Jensen. (2003b) Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiol. Mol. Biol. Rev.* 67:303-342.

CHAPTER FOUR -

PROBING THE BIOSYNTHESIS OF SCYTONEMIN THROUGH STABLE ISOTOPE INCUBATION STUDIES

Abstract

Scytonemin is a dimeric indolic-phenolic pigment found in the sheath of many cyanobacteria. This pigment absorbs UV radiation protecting the cyanobacterial cells from the harmful exposure to this radiation. Scytonemin's unique chemical structure led to an interest in its unpredictable biosynthesis. This study reports the incorporation of both tyrosine and tryptophan into the scytonemin chemical structure, and provides *in vivo* data supporting the tryptophan origin of the ketone carbon involved in the condensation of the two biosynthetic precursors. This study also reports on the novel use of a small-scale, MALDI-TOF mass spectrometry technique to monitor the near-real-time incorporation of isotopically labeled tyrosine during scytonemin biosynthesis.

Introduction

Scytonemin, a dimeric indolic-phenolic pigment, is found in the sheath of many species of cyanobacteria and possesses powerful ultraviolet (UV) radiation absorbing properties (Garcia-Pichel and Castenholz, 1991; Proteau *et al*., 1993). Its *in vivo* absorption in the UV-A range (λ_{max}=370nm) allows cyanobacteria to inhabit environments with high light conditions without negatively impacting photosynthesis (Dillon *et al*., 2002). Scytonemin's interesting chemical structure and ecological role as a sunscreen is complemented by its potentially valuable pharmaceutical role as a modulator of cell cycle control and inflammation. In 2002, Stevenson *et al*. identified this pigment as a reversible inhibitor of polo-like kinase 1 (PLK1), a serine/threonine kinase functioning at the G_2-M transition involved in controlling cell entry into mitosis (Stevenson *et al*., 2002a; **Figure 4.1**). Scytonemin was also shown to have anti-inflammatory properties when tested in a mouse ear edema assay. The dual inhibitory activities of scytonemin make it a potential pharmacophore for disorders featuring cell proliferation and inflammation in their pathology, including psoriasis, rheumatoid arthritis and asthma (Stevenson *et al*., 2002b).

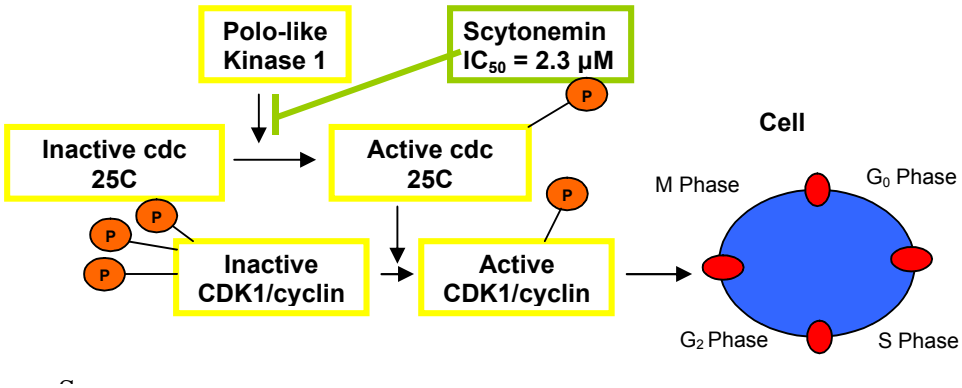


Figure 4.1: Diagram outlining protein kinase driven cell signaling involved in the G_2-M cell cycle transition. Recreated from reference Stevenson *et al*., 2002a.

Scytonemin's unique dimeric structure, ecological importance and novel pharmacological activity led to an interest in its biosynthesis (Sorrels, *et al*., 2005). The molecular scaffold of scytonemin is predicted to be biosynthesized from indolic and phenolic amino acid derived precursors (Proteau *et al*., 1993). In 2007, Soule *et al*. identified a biosynthetic gene cluster involved in the biosynthesis of scytonemin in *Nostoc punctiforme* ATCC 29133 using *in vivo* transposon mutagenesis (**Figure 4.2**). The mutation, found to inhibit scytonemin production, was located in a hypothetical protein (Np1273) associated with a gene cluster containing genes predicted to be involved in aromatic amino acid biosynthesis (Soule *et al*., 2007). The association of these genes as a cluster was later supported by evidence for increased transcriptional expression of each gene in the cluster upon exposure of the cyanobacterium to UV radiation (Sorrels *et al*., 2009).

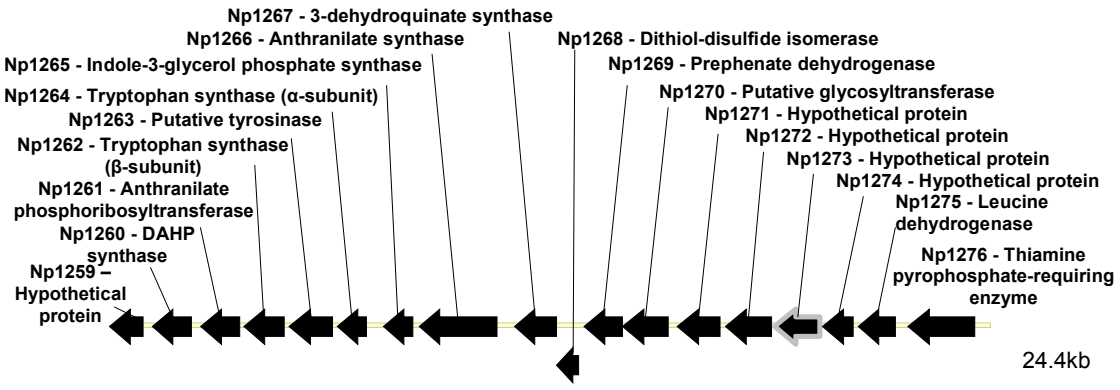


Figure 4.2: The biosynthetic gene cluster for scytonemin production as identified through random insertion of a transposon at Np1273 outlined in grey. Figure derived from Soule *et al*., 2007.

The identification of the biosynthetic gene cluster involved in the production of scytonemin led to a better understanding of the enzymatic mechanisms available for biosynthesis. The indolic precursor is predicted to be derived from the tryptophan branch of aromatic amino acid biosynthesis, while a hydroxyphenylpyruvate (HPP) precursor is created from prephenate using prephenate dehydrogenase (Np1269). Tryptophan was expected to be modified to form indole-3-acetic acid (IAA) during this process. A critical step in the biosynthesis of scytonemin is predicted to be the facilitation of a decarboxylation reaction by a thiamine pyrophosphate requiring enzyme (Np1276) that catalyzes bond formation between HPP and IAA (Sorrels *et al*., 2007; **Figure 4.3**). Recently, a recombinant Np1275 was shown to be involved in the oxidative deamination of tryptophan to form indole-3-pyruvic acid (IPA) instead of IAA. When IPA and HPP were incubated with Np1276 and cofactors, this recombinant enzyme was shown to initiate a selective acyloin reaction predicted to

form the precursor to the scytonemin monomer (Balskus and Walsh, 2008; **Figure 4.4**).

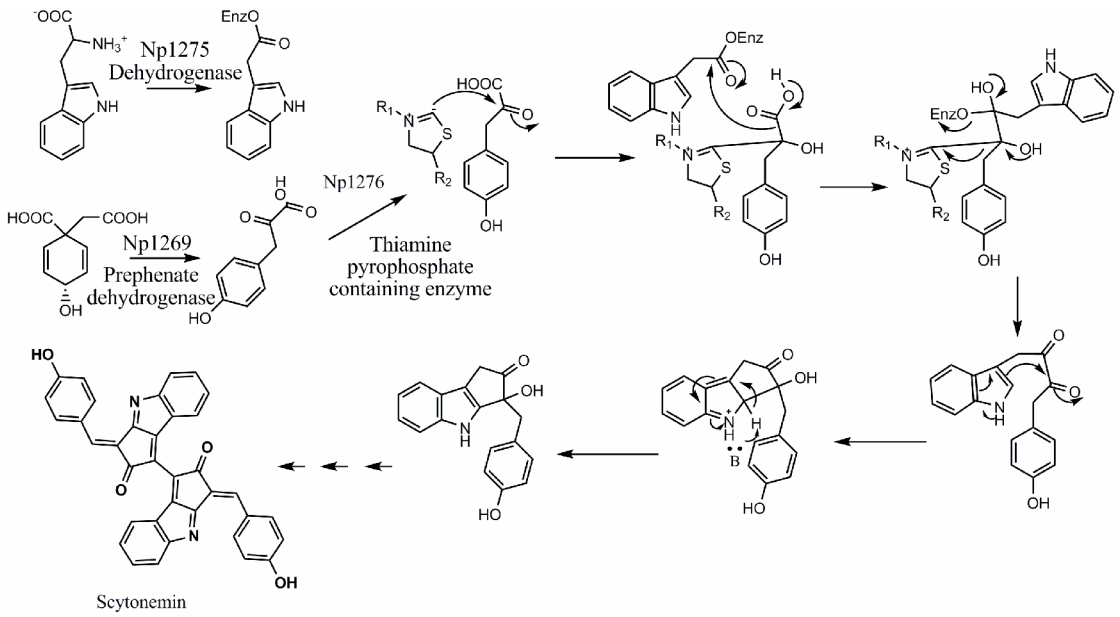


Figure 4.3: Predicted mechanism for scytonemin biosynthesis based on tryptophan and prephenate precursors. Scheme shows enzymes predicted to catalyze the conversion of these precursors to an activated form of indole-3-acetic acid and 4-hydroxyphenylpyruvate followed by condensation leading to the monomeric scytonemin moiety (Sorrels *et al*., 2007).

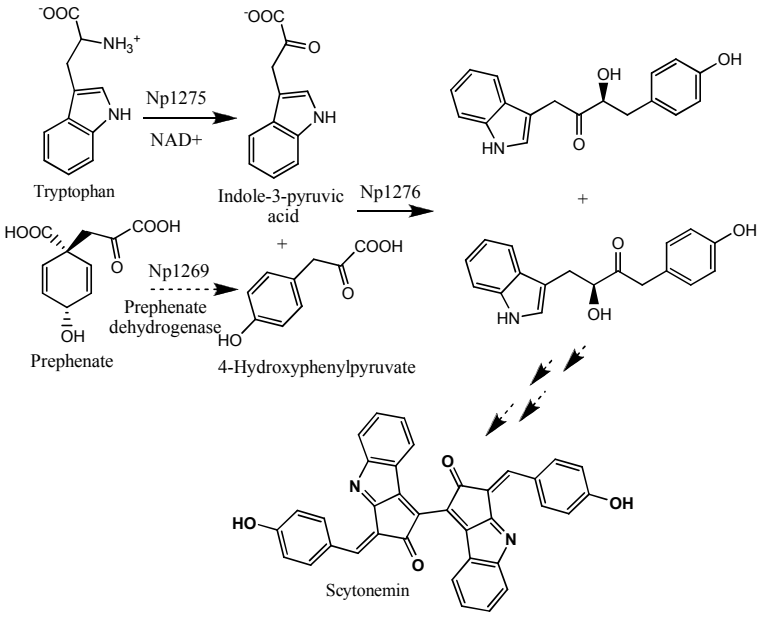


Figure 4.4: Summary of the results for enzymatic analyses of recombinant Np1275 and Np1276. Np1275 was shown to convert tryptophan to indole-3-pyruvic acid through oxidative deamination. Np1276 was shown to catalyze the formation of isomeric acyloins from indole-3-pyruvic acid and 4-hydroxyphenylpyruvic acid.

Despite the suggestive nature of this biosynthetic mechanism, *in vivo* data supporting the incorporation of these aromatic amino acid derived substrates as well as a clear understanding of the biogenesis of the carbon framework for scytonemin is needed and could provide further support for this mechanism. In this study, we report the incorporation of both tyrosine and tryptophan into the scytonemin chemical structure while showing that the ketone carbon involved in the condensation of the two biosynthetic precursors to form the scytonemin monomer is derived from the tryptophan subunit rather than the tyrosine derived subunit (**Figure 4.5**). We also use

a small-scale, MALDI-TOF mass spectrometry technique to monitor the incorporation of isotopically labeled tyrosine during scytonemin biosynthesis.

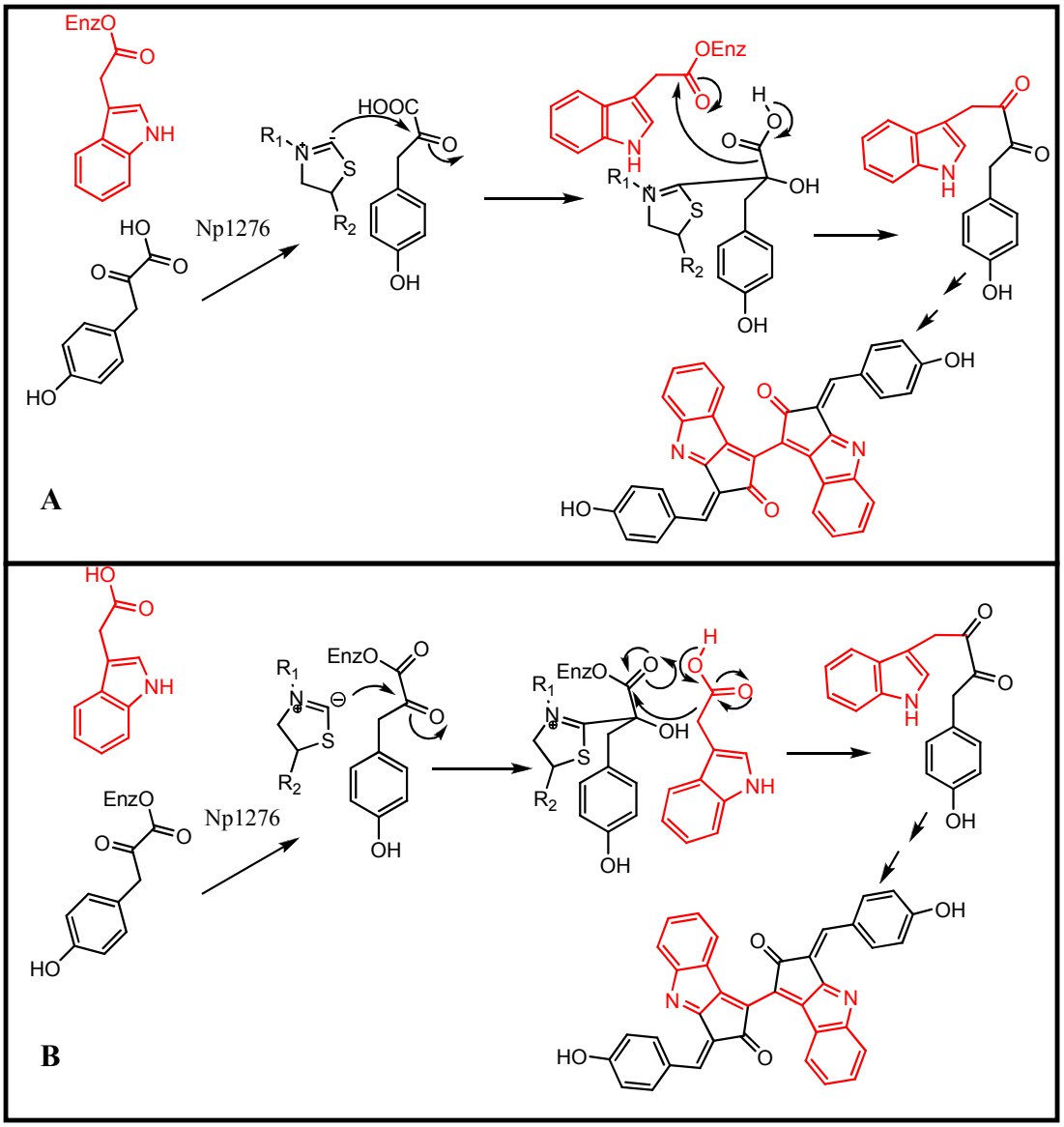


Figure 4.5: Diagram of two proposed mechanisms for the formation of a diketone precursor in scytonemin biosynthesis. A) ketone carbon of scytonemin is derived from indole-3-acetic acid (IAA), B) ketone carbon of scytonemin is derived from 4-hydroxyphenylpyruvate (HPP). IAA is outlined in red and HPP is outlined in black.

Materials & Methods

Cyanobacterial Strains and Culture Techniques

The cyanobacterium *Nostoc punctiforme* ATCC 29133 and *Nostoc* sp. ATCC 27893 (PCC 7120) were obtained from the American Type Culture Collection (ATCC). The following cyanobacteria were obtained from Carolina Biological Supply: *Gloeocapsa* sp., *Lyngbya* sp., *Oscillatoria tenius*, *Microcystis aeruginosa*, *Cylindrospermum* sp., *Fischerella musicola*, *Gloeotrichia* sp., *Anabaena inequalis*, and *Tolypothrix distorta*. The cultures were maintained in a unialgal condition in liquid BG-11 freshwater media at 20^{o}C or 29^{o}C under a light intensity of approximately 19 μmol m^{-2} s^{-1} and a light/dark cycle of 16 h/8 h. Scytonemin production was induced by exposure to 0.64 mW/cm^2 UV-A radiation (λ_{max} = 365nm) at 29^{o}C.

Instrumentation

Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance DRX300. Spectra were referenced to a residual pyridine solvent signal with resonance at δ_H 8.74. Atmospheric Chemical Ionization (APCI) mass spectra were obtained on a Thermo Finnigan LCQ Advantage mass spectrometer. Matrix Assisted Laser Desorption-Time of Flight (MALDI-TOF) spectra were obtained on a Bruker Microflex MALDI-TOF mass spectrometer using a α-cyano-4-hydroxycinnamic acid matrix coating.

L-Tyrosine-3,5-d_2 Stable Isotope Feeding Experiments

N. punctiforme ATCC 29133 was grown for approximately 75 days prior to exposure to UV-A radiation for 6 days. L-Tyrosine-3,5-d_2 (Tyr-d2; 500 mg) was added to approximately 3L of culture on days 2 and 4, and harvested on day 6 (10.91 g wet wt). The harvested mass was extracted repeatedly with methanol followed by extraction with ethyl acetate to yield 24.3 mg crude scytonemin. The crude scytonemin was repeatedly purified with methanol and hexane to yield 5.4 mg of pure isotopically labeled scytonemin.

T. distorta grown for approximately 50 days was exposed to UV-A radiation for 3 days. Tyr-d_2 (3 mg) was added to approximately 10 mL of culture at 48 hours, and harvested at 72 hours (494.1 mg wet wt). A tuft of *T. distorta* was prepared for MALDI-TOF analysis by dissolving in the MALDI matrix. The remainder was extracted repeatedly with methanol followed by extraction with ethyl acetate to yield 0.6 mg of crude labeled scytonemin.

1-$^{13}C_1$ Tyrosine Stable Isotope Feeding Experiment

N. punctiforme ATCC 29133 culture grown for approximately 75 days was exposed to UV-A radiation for 6 days. 1-$^{13}C_1$ L-Tyrosine (139 mg) and L-tyrosine-3,5-d_2 (36 mg) were added to approximately 3L of culture on days 2 and 4, and harvested on day 6 (9.89 g wet wt). The harvested mass was extracted repeatedly with methanol followed by extraction with ethyl acetate to yield 215.13 mg crude scytonemin. The crude scytonemin was repeatedly purified with methanol and hexane to yield 65.9 mg of pure labeled scytonemin.

L-Tryptophan-indole-d_5 Stable Isotope Feeding Experiment

T. distorta grown for approximately 50 days was exposed to UV-A radiation for 3 days. L-Tryptophan-indole-d_5 (Trp-d_5; 4 mg) was added to approximately 10 mL of culture at 48 hours, and harvested at 72 hours (415.9 mg wet wt). A tuft of *T. distorta* was prepared for MALDI-TOF analysis by dissolving in MALDI matrix. The remainder was extracted repeatedly with methanol followed by extraction with ethyl acetate to yield 1.7 mg of crude labeled scytonemin.

Tyrosine-d_2 and Tryptophan-d_5 Stable Isotope Feeding Experiment

T. distorta grown for approximately 50 days was exposed to UV-A radiation for 3 days. Tyr-d_2 (2 mg) and Trp-d_5 (2 mg) were added to approximately 10 mL of culture at 48 hours, and harvested at 72 hours (463.8 mg wet wt). A tuft of *T. distorta* was prepared for MALDI-TOF analysis by dissolving in MALDI matrix. The remainder was extracted repeatedly with methanol followed by extraction with ethyl acetate to yield 0.9 mg of crude labeled scytonemin.

U-$^{13}C_9$, ^{15}N Tyrosine Stable Isotope Feeding Experiment

T. distorta grown for approximately 65 days was exposed to UV-A radiation for 8 days. U-$^{13}C_9$, ^{15}N-L-Tyrosine (6 mg) was added to approximately 75 mL of culture at 48 hours, and harvested at 8 days (3.1 g wet wt). During the first 6 days (144 hours), a small tuft of *T. distorta* was removed from the culture every 8 hours, flash frozen and stored at -20^{o}C for future MALDI-TOF analysis. Between 48 hours and 64 hours, a small tuft of *T. distorta* was removed from the culture every hour,

flash frozen and stored at -20^{o}C for future MALDI-TOF analysis. These samples were prepared for MALDI-TOF analysis by dissolving in MALDI matrix.

MALDI-TOF data was analyzed using Clinprotools software (Bruker) to identify the intensity of the scytonemin mass peak and all isotopic peaks associated with natural abundance of stable isotope incorporation and incorporation of the isotopically labeled substrate. Percent isotopic incorporation of single and double isotopically labeled scytonemin was then calculated using previously established methods (Biemann, 1961). The rate of incorporation of U-$^{13}C_9$, ^{15}N-tyrosine into scytonemin was determined by averaging the percent intensity of the mass peak at *m/z* 554 based on the monoisotopic peak at *m/z* 546. This data was plotted using Excel (Microsoft).

Results and Discussion

Scytonemin's unique molecular scaffold containing both an indolic and phenolic moiety suggested that these units may be derived from tryptophan and tyrosine respectively. The incorporation of these amino acid derived subunits during scytonemin biosynthesis can be validated using stable isotope enrichment studies. Preliminary analyses incubating *N. punctiforme* ATCC 29133 with unlabeled L-tryptophan suggested that this amino acid was not efficiently transported into the cell. As a result, the media became discolored compared to control cultures during the course of UV radiation. Due to the results of this preliminary study, tyrosine was

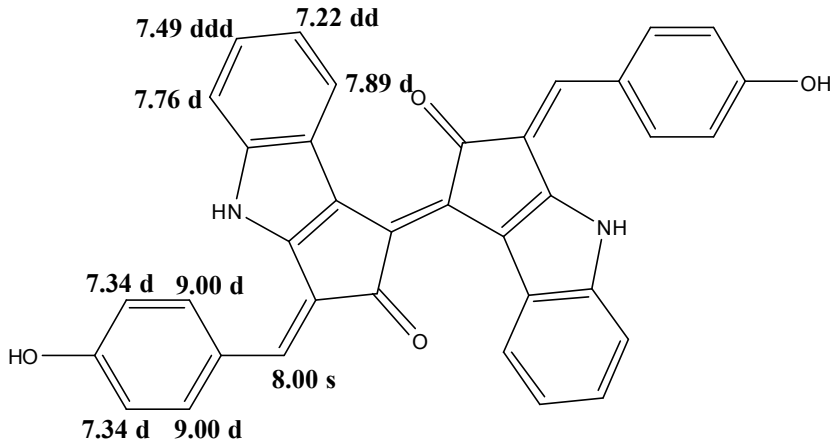
chosen as the initial substrate to begin stable isotope enrichment studies in order to probe the biogenetic origin of the carbons found in the scytonemin molecular scaffold.

L-tyrosine-3,5-d_2 (Tyr-d_2), labeled at the meta position of the phenol ring with deuterium, was incubated with *N. punctiforme* ATCC 29133 filaments while being exposed to UV-A radiation. Analysis of the purified isotopically labeled scytonemin compared to control unlabeled scytonemin by ^{1}H NMR (**Figure 4.6**) revealed a decrease in the intensity of the doublet signal at 7.33 ppm when compared to the doublet at 7.76 ppm. These results indicate that deuterium is partially replacing hydrogen at the meta position of the phenol ring in scytonemin. This incorporation is further supported by the appearance of an apparent triplet at 9.0 ppm. The apparent triplet forms when the hydrogen at 7.33 ppm is replaced by deuterium, which eliminates the coupling between the two protons in the ortho and meta positions of the phenol ring giving rise to a singlet. The overlap of a singlet and doublet thus forms an apparent triplet. The incorporation of this isotopically labeled tyrosine substrate reveals that the phenolic portion of the scytonemin scaffold is unequivocally derived from the aromatic amino acid tyrosine.

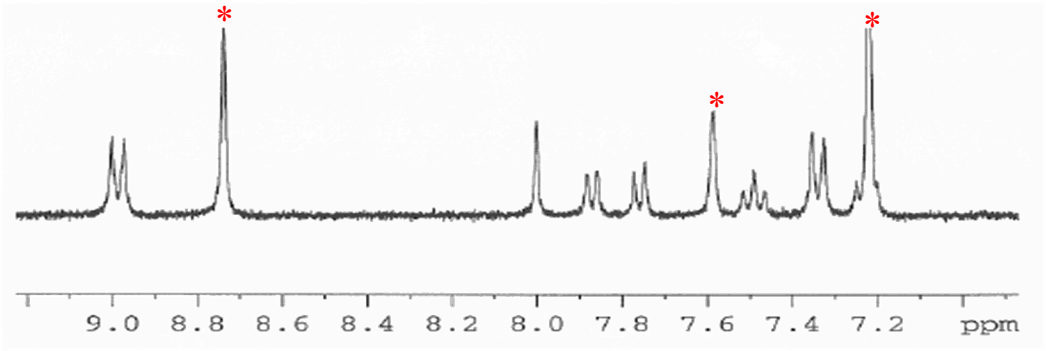
Tyrosine has recently been shown to act as a competitive inhibitor of prephenate dehydrogenase through a predicted interaction between the amino group of the tyrosine and the main chain carbonyl of Thr-152 in the active site of the enzyme in the bacterium *Aquifex aeolicus* (Sun et al., 2009). Interestingly, the scytonemin biosynthetic gene cluster contains a prephenate dehydrogenase that is predicted to catalyze the decarboxylation of prephenate to create the scytonemin precursor, 4-

hydroxyphenylpyruvate (Balskus and Walsh, 2008; Sorrels *et al*., 2007). The ability of tyrosine to act as a competitive inhibitor of this enzyme would suggest that when isotopically labeled tyrosine is present in the culture media, the activity of the prephenate dehydrogenase is reduced. The use of exogenous tyrosine by *N. punctiforme* ATCC 29133 in the biosynthesis of scytonemin indicates that enzymes not associated with the biosynthetic gene cluster, such as a tyrosine transaminase, may be utilized to produce the phenolic substrate needed for biosynthesis.

A



B



C

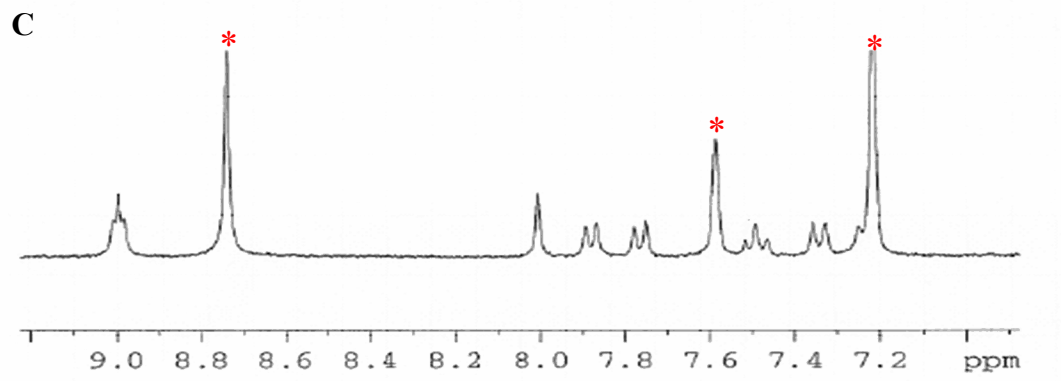


Figure 4.6: A) Structure of scytonemin labeled with reported ^{1}H resonances when in pyridine solvent (Proteau *et al*., 1993), B) ^{1}H NMR of unlabeled scytonemin C) ^{1}H NMR of scytonemin after incorporation of L-tyrosine-3,5-d_2. Pyridine solvent peaks marked with red asterisk.

Mass spectrometry is also a powerful tool for the analysis of isotopically labeled secondary metabolites. Scytonemin has a molecular formula of $C_{36}H_{20}N_2O_4$ and is reported to have a mass to charge ratio (*m/z*) of 544. Atmospheric Chemical Ionization mass spectrometry (APCI-MS) in the positive ionization mode, reveals the molecular ion (M+H) at *m/z* 545 for monoisotopic scytonemin as seen in **figure 4.7a**. The smaller peaks following the monoisotopic peak represent the incorporation of ^{13}C into the molecule due to its natural abundance of 1.11% (Crews et al., 1998). When the isotopically labeled scytonemin is analyzed using the same mass spectrometry conditions, the addition of two and four mass units are clearly present. This increase in mass represents the incorporation of one or two of the tyr-d_2 molecules containing the deuterium labels at the meta positions of the phenol ring (**Figure 4.7b**). Based on these results the level of stable isotope incorporation is calculated to be approximately 35% deuterated scytonemin (Biemann, 1961).

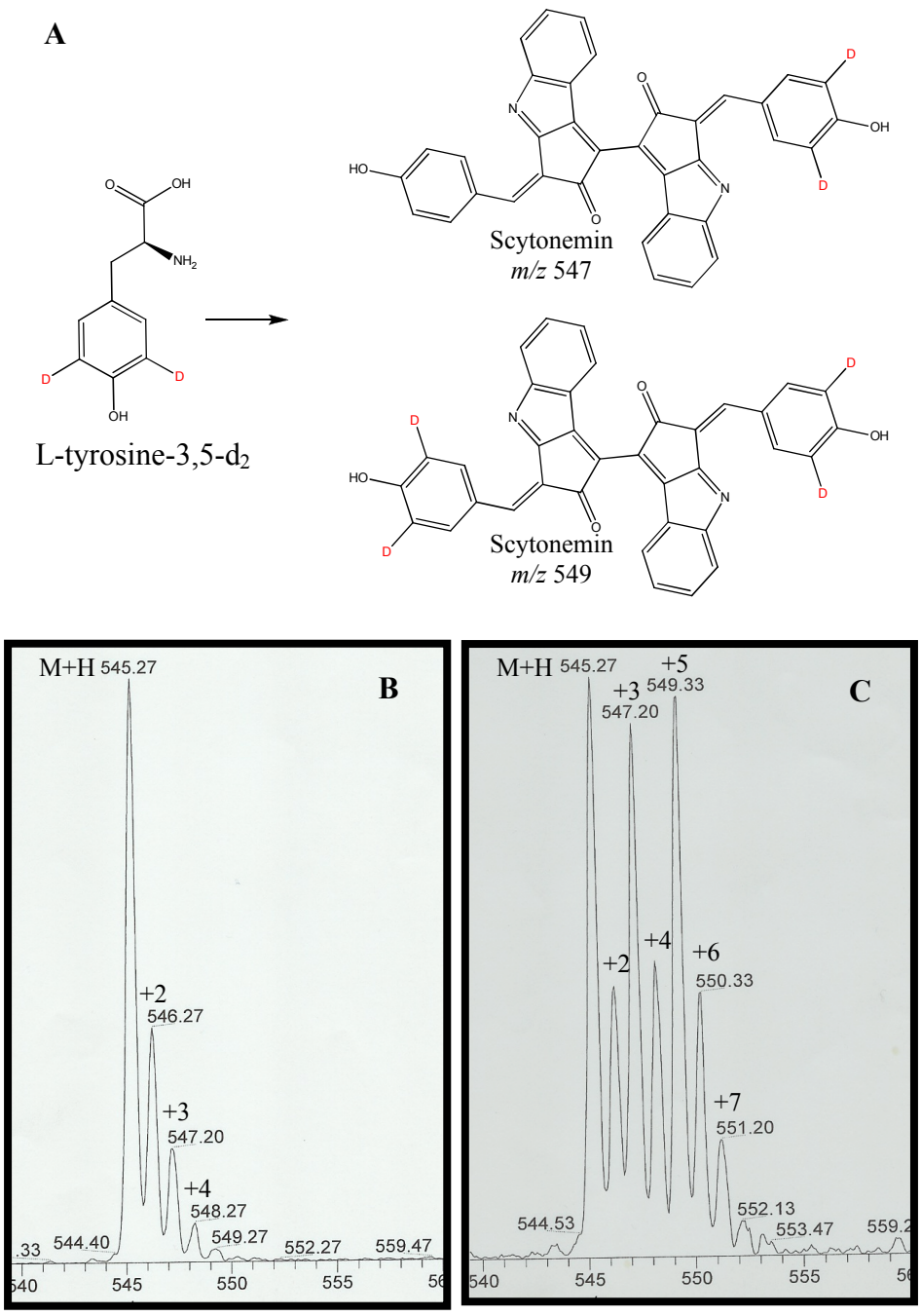


Figure 4.7: Positive ionization mode APCI-MS data; A) chemical structure of the isotopically labeled precursor and the resulting scytonemin products, B) Standard scytonemin sample C) Scytonemin enriched with approximately 35% L-tyrosine-3,5-d_2.

In 2005, cultures of *N. punctiforme* ATCC 29133 were transported south about 1,000 miles from Corvallis, OR (Oregon State University) to La Jolla, CA (University of California, San Diego). The transition of these cultures to a new environment led to slower growth and limited scytonemin production, even after exposure to UV-A radiation. In attempts to regain the production of higher quantities of scytonemin, these cultures were exposed to various environmental conditions including different media, increased and decreased levels of micronutrients, different light sources, and different temperatures; however, the production of scytonemin still only occurred in minimal amounts and with unpredictable frequency. Although scytonemin production is known to be induced by UV radiation, it has also been shown in *Calothrix* sp. to be dependent on other environmental conditions. In this study by Dillon and Castenholz, two strains of *Calothrix* sp., both capable of producing scytonemin, were taken from two springs within the same region. The level of scytonemin production in the individual strains were influenced by the aquatic environment suggesting that scytonemin production is not based solely on the availability of UV-A radiation (Dillon and Castenholz, 2003). During our studies, it is possible that a small environmental parameter differing between Oregon and California culture laboratories played a role in allowing for the production of higher levels of scytonemin in *N. punctiforme* ATCC 29133 while in Oregon.

In order to complete the stable isotope enrichment studies and determine the biogenesis of the carbons in scytonemin, another scytonemin producing cyanobacterial strain was needed. Ten freshwater species of cyanobacteria were placed under UV

radiation to identify their ability to produce scytonemin. Of these species, only 5 survived the UV radiation exposure including *Nostoc* sp. PCC7120, *Fischerella musicola*, *Lyngbya* sp., *Oscillatoria tenius*, and *Tolypothrix distorta*. These five species and a *N. punctiforme* ATCC 29133 culture were analyzed for the presence of the scytonemin using MALDI-TOF mass spectrometry (**Figure 4.8**). *T. distorta* was identified as a scytonemin producing cyanobacterial species due to the monisotopic mass located at *m/z* 546 (**Figure 4.8H**) and an isotopic ratio pattern matching the purified scytonemin analyzed as a control (**Figure 4.8C**). *T. distorta* was also shown to produce scytonemin in quantities reliable for analytical analyses; therefore, this species was chosen for the remainder of the stable isotope enrichment studies.

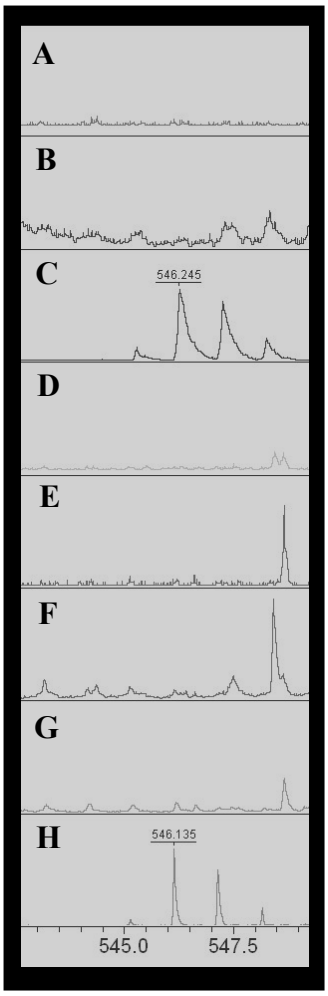


Figure 4.8: MALDI-TOF mass spectrometry results for the screening of cyanobacterial cultures for scytonemin production. A) Matrix control, B) *N. punctiforme* ATCC 29133, C) Purified scytonemin, D) *Nostoc* sp. PCC7120, E) *F. musicola*, F) *Lyngbya* sp., G) *O. tenius*, H) *T. distorta*. Mass peak at *m/z* 546 represents the scytonemin monoisotopic species.

The use of MALDI-TOF mass spectrometry for small-scale analyses of UV-induced *T. distorta* provided a quick and cost effective tool for the analysis of stable isotope enrichment studies in this organism. This approach is validated by comparing the results of a small scale *T. distorta* incubation study using tyr-d_2 with those results previously obtained from *N. punctiforme* ATCC 29133 at a larger scale. Using only a single tuft (approximately 7 mg wet wt.) from the *T. distorta* culture after exposure to UV radiation, the incorporation of Tyr-d_2 gives rise to the same isotopic mass profile as previously seen in *N. punctiforme* ATCC 29133 after a 48 hour incubation (**Figure 4.9**). This profile shows an isotopic mass peak at [M+2] *m/z* 548 and [M+4] *m/z* 550 representing the incorporation of either one or two of the isotopically labeled tyrosines into the scytonemin scaffold, respectively. The use of the MALDI-TOF technique to analyze the incorporation of the isotopically labeled tyr-d_2 required a very small amount of label (<10 mg), only minimal amounts of the cultured organism, and allowed for the entire experiment to be completed in one quarter the time needed for the same incubation study in *N. punctiforme* ATCC 29133, making this technique a valuable tool for exploring biosynthesis through the analysis of isotopically enriched secondary metabolites.

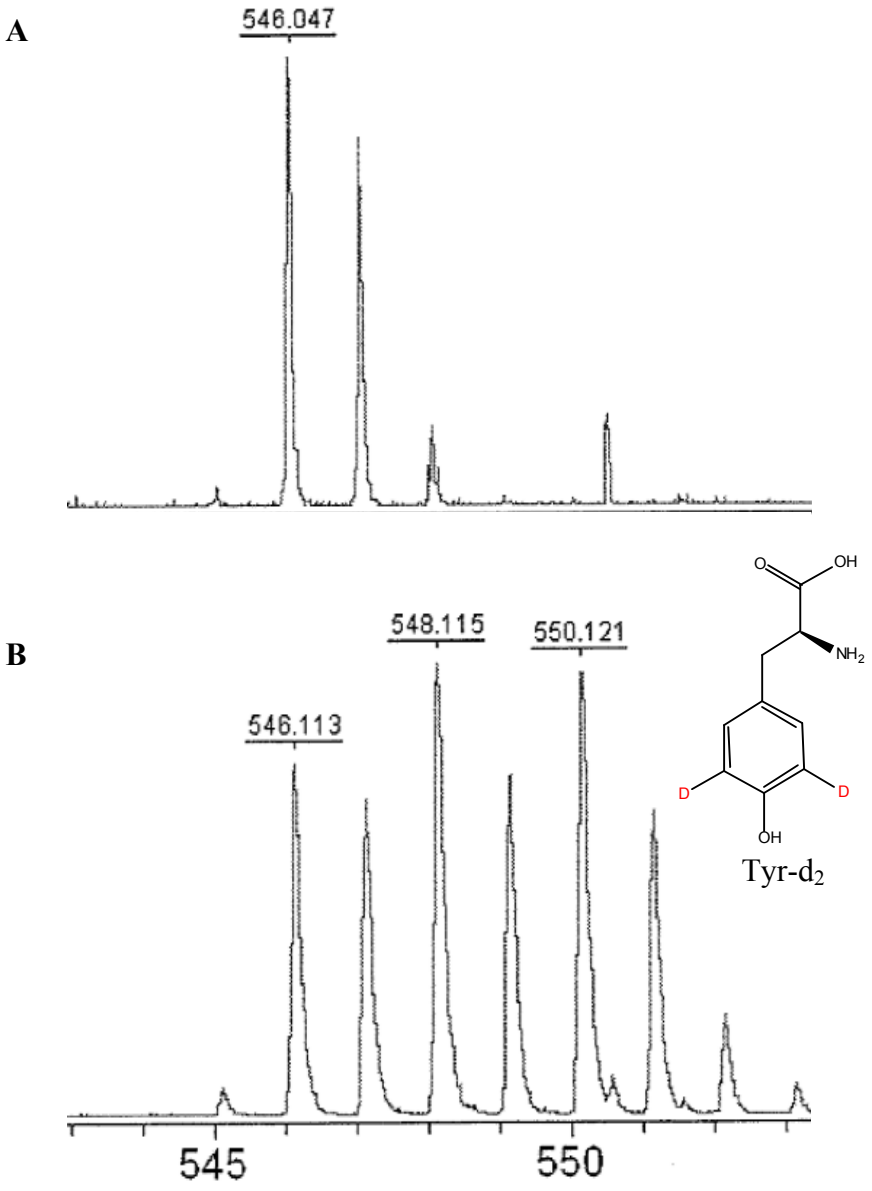


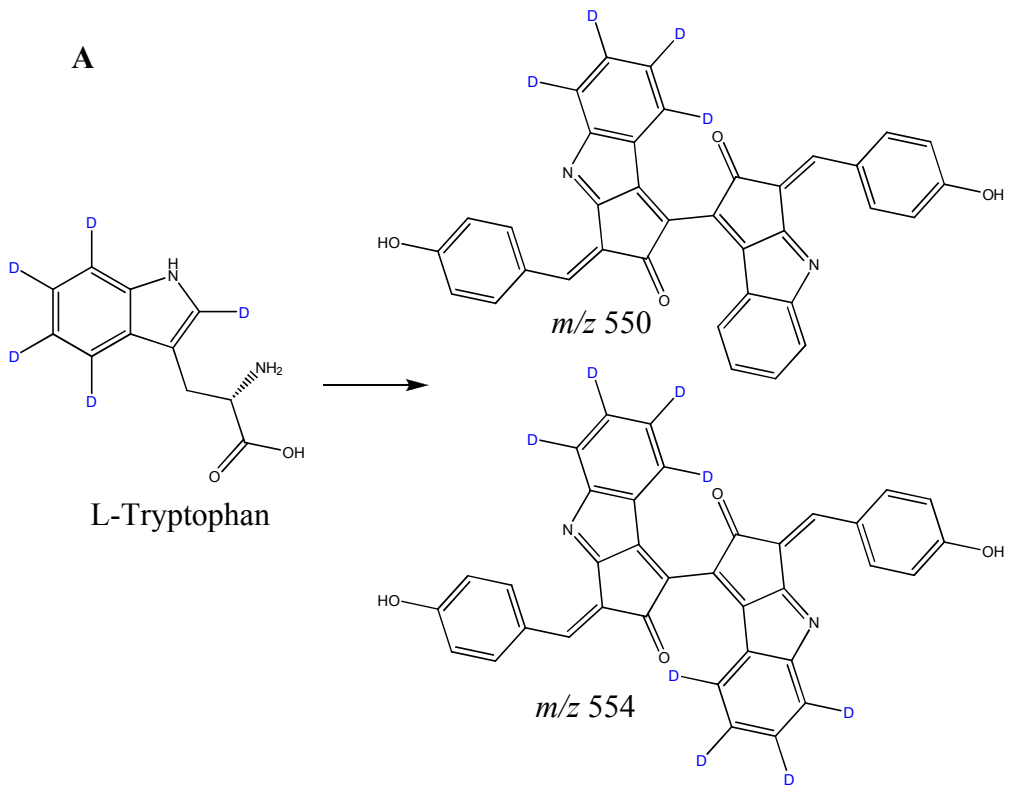
Figure 4.9: MALDI-TOF spectra of *T. distorta* after incubation with L-tyrosine-3,5-d_2 during UV radiation treatment. A) Spectra of control *T. distorta* without tyr-d_2 incubation, B) spectra of *T. distorta* after incubation with isotopically labeled tyrosine. Structure of tyr-d_2 shown in B.

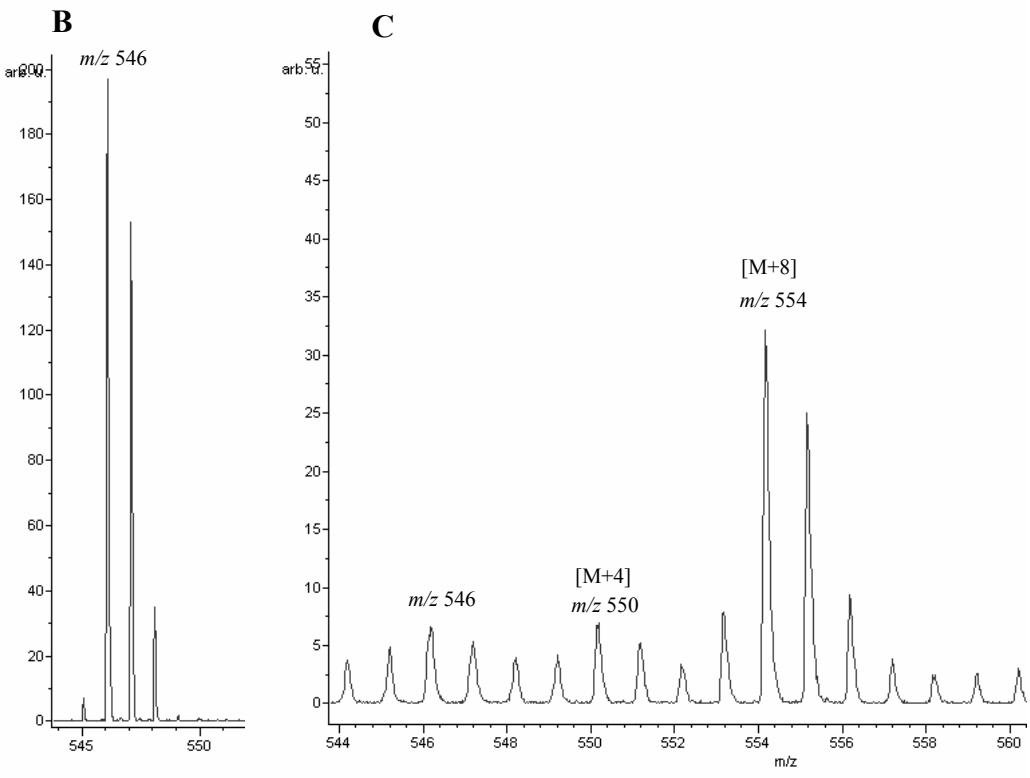
The sensitivity of MALDI-TOF and the short length of incubation time prior to analysis allowed us to overcome the initial problems with using isotopically labeled tryptophan as a substrate for enrichment of scytonemin. An initial experiment using L-tryptophan-indole-d_5 as the incubated substrate resulted in an intense [M+8] peak at *m/z* 554 and an [M+4] peak at *m/z* 550 (**Figure 4.10**). The [M+8] peak represents the incorporation of two isotopically label tryptophans into scytonemin, while the [M+4] peak represents the incorporation of only a single isotopically labeled tryptophan. The incorporation of only four of the five available deuteriums per tryptophan subunit is not surprising as the cyclization of the intermediate to form the scytonemin monomer is expected to remove the hydrogen located at the 2 position of the indole ring resulting in a five membered ring (**Figure 4.3**). Therefore, the enrichment of scytonemin by four and eight mass units is consistent with tryptophan being used during the biosynthesis of this metabolite.

Due to the energetic cost of making aromatic amino acids, tryptophan, like tyrosine, also acts as an allosteric inhibitor towards its own biosynthesis, specifically by inhibiting the enzyme anthranilate synthase (Merino *et al*., 2008; Xie *et al*., 2003). The presence of an anthranilate synthase in the biosynthetic gene cluster for scytonemin and the evolutionary conservation of enzymes found in the tryptophan operon suggests that this enzyme may also be inhibited by the presence of tryptophan (Xie et al., 2003). Interestingly, the isotope enrichment study using a tryptophan precursor suggests a limited endogenous biosynthesis of tryptophan in the presence of exogenous tryptophan. This is indicated by the intensity of the [M+8] isotopic peak at

m/z 554. The monoisotopic peak and the [M+4] peak have very low intensities when compared to the [M+8] peak, indicating that most of the scytonemin produced during this time period is created using the isotopically labeled tryptophan.

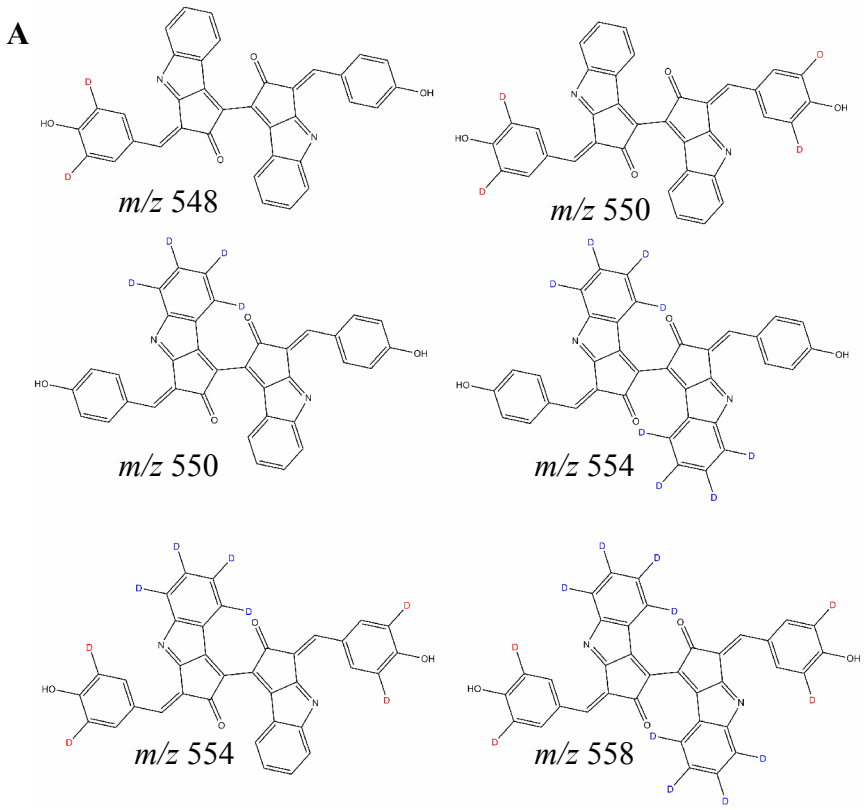
Figure 4.10: MALDI-TOF spectra of *T. distorta* after incubation with L-tryptophan-indole-d_5 during UV stimulation. A) Chemical structures of precursor and resulting scytonemin products, B) control spectrum of *T. distorta* showing the scytonemin monoisotopic peak at *m/z* 546; C) spectrum of *T. distorta* after incubation with L-tryptophan-indole-d_5 showing a small amount of scytonemin with the incorporation of one tryptophan (*m/z* 550) and a significant amount of scytonemin with the incorporation of two isotope enriched tryptophan subunits (*m/z* 554).

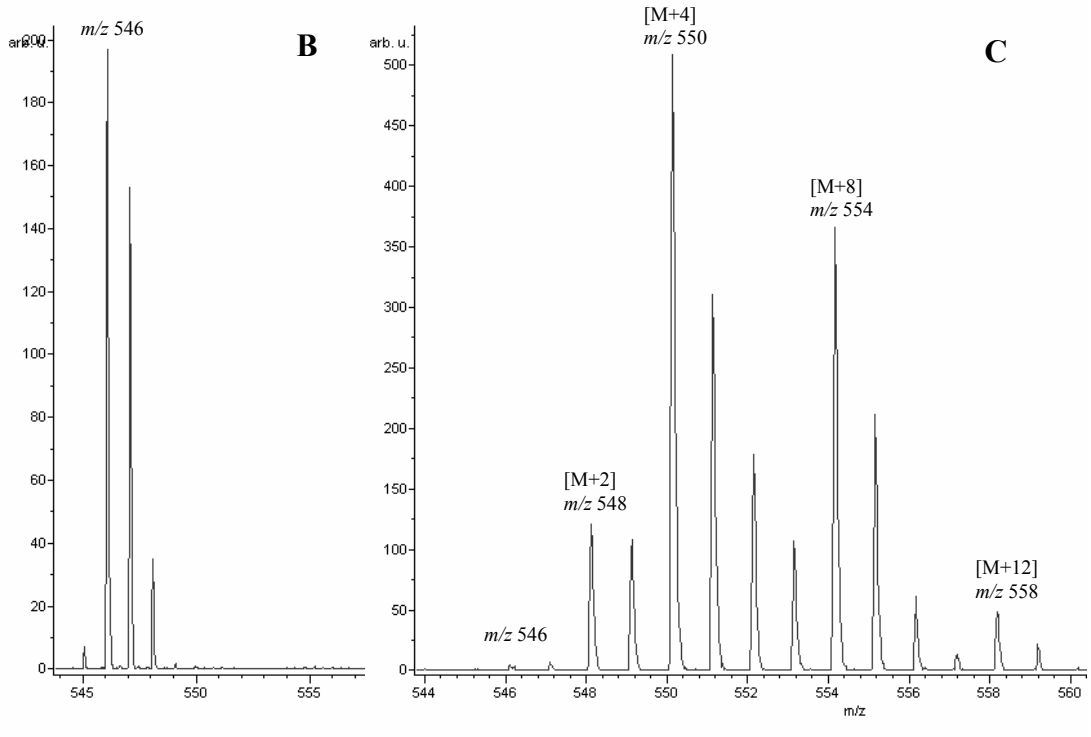

A
L-Tryptophan
m/z 550
m/z 554


B
m/z 546
C
[M+8]
m/z 554
[M+4]
m/z 550
m/z 546
m/z

L-tryptophan-indole-d_5 and L-tyrosine-3,5-d_2 were also shown to be incorporated into the same scytonemin molecule. The MALDI-TOF results show large mass intensities for the [M+4] peak at *m/z* 550, the [M+8] peak at *m/z* 554, and the [M+12] peak at *m/z* 558 (**Figure 4.11**). The [M+4] peak at *m/z* 550 represents either the incorporation of one deuterated tryptophan subunit or two deuterated tyrosine subunits, and the [M+8] peak at *m/z* 554 represents the incorporation of either two deuterated tryptophan subunits or one deuterated tryptophan and two deuterated tyrosine subunits into the scytonemin molecular skeleton. The critical isotopic mass for determining the incorporation of both tryptophan and tyrosine into the same molecule is found at the [M+12] peak at *m/z* 558. This mass results from the incorporation of both two deuterated tryptophan subunits each giving rise to four deuteriums (8 total) and two deuterated tyrosine subunits each giving rise to two deuteriums (4 total). These results support tyrosine and tryptophan as substrates for scytonemin biosynthesis and demonstrate the utilization of MALDI-TOF for incubation studies using multiple substrates.

Figure 4.11: MALDI-TOF spectra of *T. distorta* after incubation with L-tryptophan-indole-d_5 and L-tyrosine-d_2 during stimulation by UV radiation. A) Chemical structures for isotopically labeled scytonemin molecules seen by MALDI-TOF, B) Control spectrum of *T. distorta* without incubation with isotopically labeled substrates showing the scytonemin monoisotopic peak at *m/z* 546, C) mass spectrum of *T. distorta* after incubation with isotopically labeled substrates showing significant isotopic peaks at [M+2] representing the incorporation of a single tyrosine, [M+4] representing the incorporation of two tyrosine subunits or a single tryptophan subunit, [M+8] representing the incorporation of two tyrosines and one tryptophan subunit or two tryptophan subunits, and [M+12] representing the incorporation of two tyrosine and two tryptophan derived subunits.


A
m/z 548
m/z 550
m/z 550
m/z 554
m/z 554
m/z 558


B
m/z 546
arb. u.
C
[M+4]
m/z 550
[M+8]
m/z 554
[M+2]
m/z 548
[M+12]
m/z 558
m/z 546
m/z

MALDI-TOF mass spectrometry is not limited to the analysis of a single time point during stable isotope enrichments studies. This technique can be expanded to gain insights into the rate of biosynthesis of secondary metabolites. Using U-$^{13}C_9$, ^{15}N tyrosine as a substrate for incubation studies in *T. distorta*, we gained a unique glimpse of the biosynthetic rate of tyrosine utilization in the formation of scytonemin throughout a 6 day enrichment study. During this study, small tufts of *T. distorta* were taken every 8 hours and analyzed using MALDI-TOF for scytonemin production. The monoisotopic peak for scytonemin at *m/z* 546 reached detectable levels after 24 hours of exposure to UV-A radiation. The isotopically labeled substrate was introduced into the culture after 48 hours of UV exposure to assure that the concentration of scytonemin had reached a detectable level. Previous studies have shown that scytonemin production in the cyanobacterium *Chroococcidiopsis* sp. increases the greatest between 48 and 96 hours (Dillon *et al*., 2002).

The parent peak [M+8] for the isotopically labeled scytonemin at *m/z* 554 reaches detectable levels at 49 hours after exposure to UV radiation, only one hour after the introduction of the isotopically labeled substrate into the culture. **Figure 4.12** shows the increase in the [M+8] peak over the first 16 hours of incubation with the isotope enriched precursor.

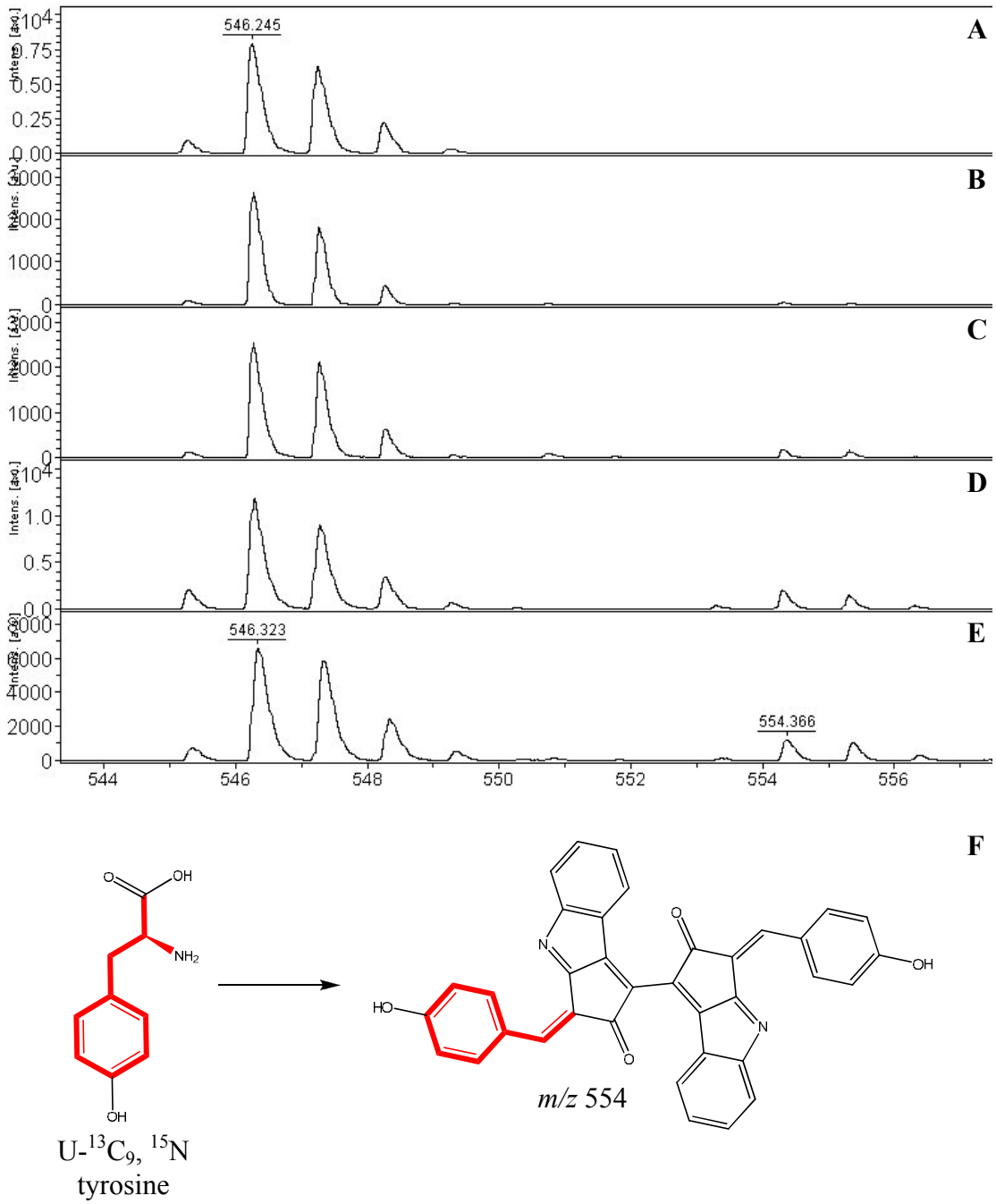


Figure 4.12: MALDI-TOF spectra showing the incorporation of U-$^{13}C_9$, ^{15}N tyrosine into scytonemin during the first 16 hours after introduction of the isotopically enriched substrate to *T. distorta* filaments under UV radiation. A) 48 hours, B) 52 hours, C) 56 hours, D) 60 hours, E) 64 hours, F) chemical structure of precursor and resulting isotopically labeled scytonemin.

By comparing the percent intensity of the [M+8] peak to the percent intensity of the [M+] peak at *m/z* 546 throughout a six day incubation, the trend in isotopically labeled tyrosine can be determined (**Figure 4.13**). This trend shows a significant increase in the [M+8] peak at *m/z* 554 over the first 36 hours that the isotopically enriched substrate was available to the filaments. After this time period, a slow decline in the intensity of the [M+8] peak indicates that the cyanobacterium had exhausted the supply of exogenous tyrosine.

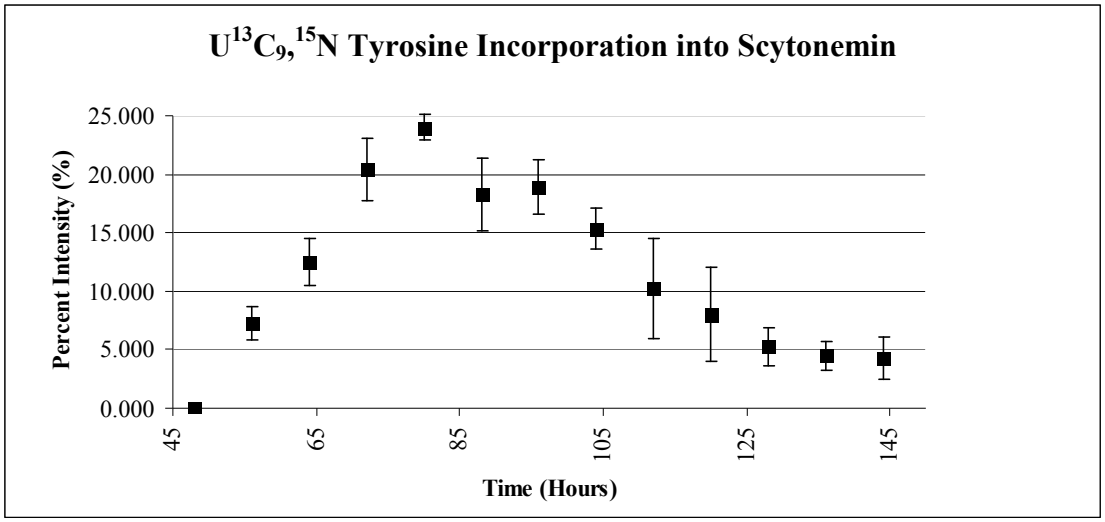
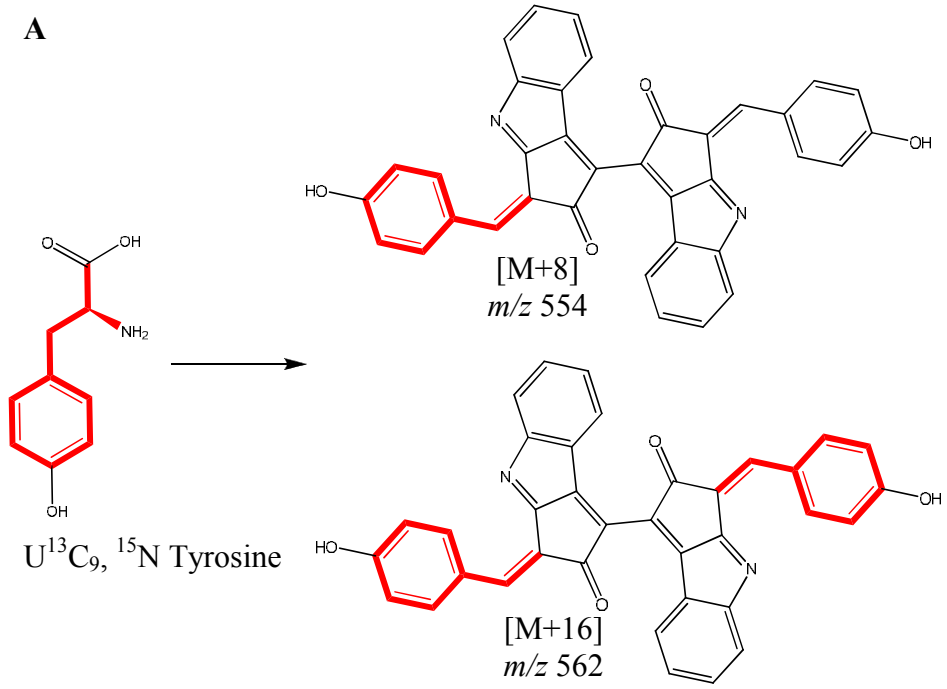


Figure 4.13: Graphical representation of the percent intensity of the parent peak of the isotopically labeled scytonemin (*m/z* 554) compared to the monoisotopic peak (*m/z* 546). An increase in the percent intensity between 48 and 84 hours represents a high utilization of the labeled substrate for scytonemin biosynthesis. The decline in percent intensity between 84 and 144 hours likely represents a decrease in the available isotope labeled substrate thus resulting in a decrease in the overall percent of isotopically labeled scytonemin. Standard error is reported from 2 biological samples, one with 3 technical replicates.

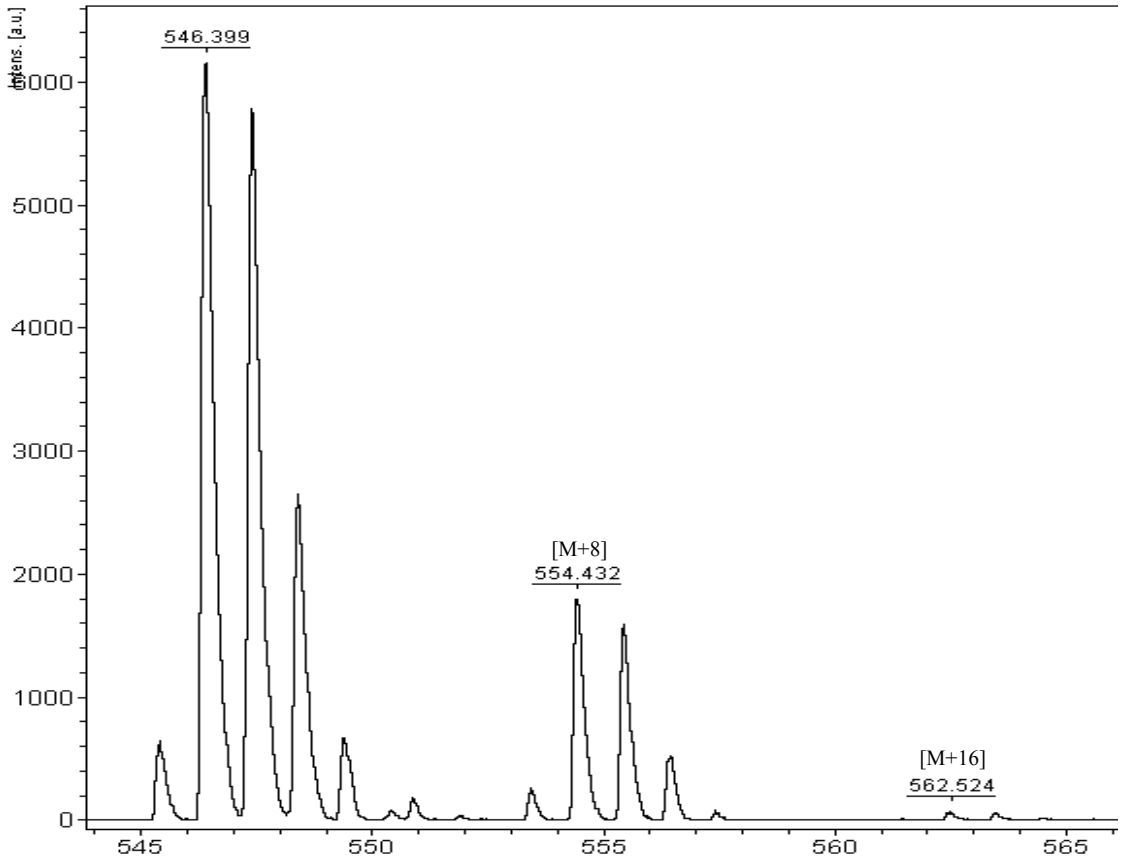
The finding that the maximum percent incorporation of isotopically labeled tyrosine into scytonemin is at 32 hours after introduction of the isotopically labeled precursor allowed us to use this time point to demonstrate the origin of the ketone carbon formed during the condensation of the indolic and phenolic subunits. **Figure 4.14** shows an [M+8] peak at *m/z* 554 and a less intense [M+16] peak at *m/z* 562. These masses represent the incorporation of either one or two isotopically labeled tyrosines into scytonemin. These results indicate that only 8 of the 10 available isotopically labeled atoms in the tyrosine are incorporated into scytonemin. The incorporation of eight carbons and the previous results showing the loss of the hydrogen at the 2 position of the indole ring in tryptophan confirm that the ketone carbon at position 2 of scytonemin is derived from the tryptophan subunit, and not the tyrosine subunit. These *in vivo* results support the previous *in vitro* proposed biosynthetic mechanism where Np1276, found in the scytonemin biosynthetic gene cluster, catalyzes the condensation of indole-3-pyruvic acid and hydroxyphenylpyruvic acid through a decarboxylation reaction to form an acyloin. This acyloin product represents the carbon skeleton required for the scytonemin monomer (Balskus and Walsh, 2008).

Figure 4.14: MALDI-TOF spectrum showing the incorporation of U-$^{13}C_9$,^{15}N tyrosine into scytonemin in *T. distorta*; A) chemical structures for the isotopically labeled scytonemin molecules in the spectrum, B) mass spectrum of the isotopically labeled scytonemin.

A


[M+8]
m/z 554
[M+16]
m/z 562
U^{13}C$_9$, ^{15}N Tyrosine

B


546.399
[M+8]
554.432
[M+16]
562.524
Intens. [a.u.]
6000
5000
4000
3000
2000
1000
0
545
550
555
560
565

The use of the MALDI-TOF method to analyze the enrichment of scytonemin over a period of time also helps explain previous results obtained in *N. punctiforme* ATCC 29133. Before transferring the filaments from Oregon to La Jolla, *N. punctiforme* ATCC 29133 was incubated with 1-$^{13}C_1$ tyrosine while being exposed to UV-A radiation at a large scale. The results of this experiment were puzzling because they showed a clear incorporation of one carbon based on mass spectrometry; however, there was no evidence for the incorporation of this isotope in any particular carbon of the scytonemin molecule based on ^{13}C NMR. During the enrichment study with *T. distorta*, an increase in the intensity of the [M+1] peak at *m/z* 547 was also apparent over time. This increased intensity was disproportionate with the expected ratio from the natural abundance of ^{13}C. The disproportionate intensity was noticeable after 24 hours of incubation with the isotopically labeled tyrosine and began to decline after 48 hours of incubation, similar to the decline seen for the intensity of the [M+8] peak at *m/z* 554 (**Figure 4.15**). The findings in *T. distorta* and *N. punctiforme* ATCC 29133 suggest that the carboxylic acid carbon of the tyrosine is being cleaved and recycled through the primary metabolic pathways of these cyanobacteria. By recycling the 1-^{13}C of tyrosine, these cyanobacteria are predicted to enrich other portions of the scytonemin molecule with an even distribution.

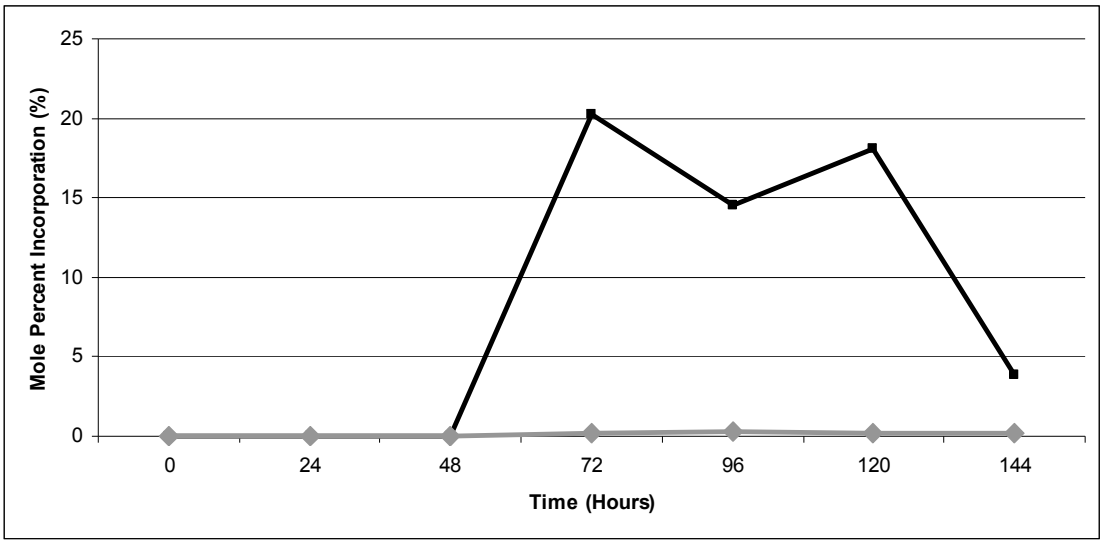


Figure 4.15: Mole percent incorporation of the [M+1] at *m/z* 547 when compared to the [M+] peak at *m/z* 546. Solid black line shows the results of the isotopically labeled *T. distorta* and the grey line is a control without incubation with the isotopic label.

This scenario explains the incorporation of one mass unit in mass spectra obtained in *N. punctiforme* ATCC 29133 fed $^{13}C_1$-1 tyrosine but with no increase in the intensity of any specific carbon atom in the ^{13}C NMR. Based on how quickly the incorporation of one mass unit is seen during scytonemin biosynthesis, it seems that during the period of active scytonemin biosynthesis (e.g. upon UV induction), a substantial portion of energy and substrates used in primary metabolism may be shunted towards this biosynthetic pathway. However, the importance of scytonemin as a last line of defense against the deadly effects of intense UV radiation may warrant

the need to direct all available resources towards the production of this protective sunscreen.

Conclusions

In this study, we showed that both tyrosine and tryptophan are substrates for the biosynthesis of scytonemin, and provide *in vivo* data to support the decarboxylation of both IPP and HPP as has been previously described as the mechanism for Np1276 (Balskus and Walsh, 2008). These findings also suggest that scytonemin biosynthesis may involve a more complicated regulation of aromatic amino acid biosynthesis and metabolism when exogenous substrates are available.

We also identified a novel use for a method designed to analyze isotopically enriched secondary metabolites in cyanobacteria by MALDI-TOF-MS. By using this method, large amounts of cultured biomass or expensive isotopically-labeled substrates were not required. Moreover, the method was rapid and gave clear results. This technique can also be used to monitor the incorporation of isotopically enriched substrates into secondary metabolites over time, thus giving insight into rates of metabolite production and turnover.

As the expense of pharmaceuticals continues to increase, the value of methods used to understand the biosynthesis of natural products is becoming ever more important. MALDI-TOF mass spectrometry coupled with isotope enrichment studies is a powerful tool being developed that will provide insights into these biosyntheses.

Acknowledgements

Chapter 4, in part, is currently being prepared for submission for publication of the material. Sorrels, C.M., Esquenazi, E., Dorrestein, P., and Gerwick, W.H. The dissertation author was the primary investigator and author of this material.

This work was supported by Sea Grant and by a National Institutes of Health predoctoral fellowship through the Training Grant in Marine Biotechnology (T32GM067550). Previous support from the Oregon Sea Grant program R/BT-40 is also acknowledged. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

I would like to acknowledge Eduardo Esquenazi for running the MALDI-TOF instrument used in these analyses and Dr. Pieter Dorrestein at the University of California, San Diego for the use of the MALDI instrument. I would also like to acknowledge Dr. Kerry McPhail at Oregon State University for assistance with NMR and Dr. Patricia Flatt for assistance with initial culturing and extraction of *N. punctiforme* ATCC 29133.

References

Balskus, E.P. and Walsh, C.T. (2008) Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J. Am. Chem. Soc*. 130: 15260-15261.

Biemann, K. (1962) The mass spectra of isotopically labeled molecules. In *Mass Spectrometry: Organic Chemical Applications*. McGraw-Hill Book Company, Inc. New York. 223-227.

Castenholz, R.W. and Garcia-Pichel, F. (2000) Cyanobacterial responses to UV-Radiation. In *The Ecology of Cyanobacteria – Their Diversity in Time and Space*. Whitton, B.A. and Potts, M. eds. Kluwer Academic Publishers. The Netherlands. 591-611.

Dillon, J.G., Tatsumi, C.M., Tandingan, P.G., and Castenholz, R.W. (2002) Effect of environmental factors on the synthesis of scytonemin, a UV-screening pigment, in a cyanobacterium (Chroococcidiopsis sp.). *Arch. Microbiol*. 177: 322-331.

Garcia-Pichel, F. and Castenholz, R.W. (1991) Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycol*. 27: 395-409.

Merino, E., Jensen, R.A., and Yanofsky, C. (2008) Evolution of bacterial trp operons and their regulation. *Cur. Opin. Microbiol*. 11: 78-86.

Proteau, P.J., Gerwick, W.H., Garcia-Pichel, F., and Castenholz, R. (1993) The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experentia*. 49: 825-829.

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2005) Biosynthesis of scytonemin: a cyanobacterial sunscreen. *American Society of Pharmacognosy*, Corvallis, OR.

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2007) Biosynthetic and chemical analyses of cyanobacterial UV-absorbing secondary metabolites. *American Society of Pharmacognosy*, Portland, ME.

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2009) Organization, evolution and expression analysis of the biosynthetic gene cluster for scytonemin, a cyanobacterial ultraviolet absorbing pigment. *Appl. Environ. Microbiol*. 75: 4861-4869.

Soule, T., Stout, V., Swingley, W.D., Meeks, J.C., and Garcia-Pichel, F. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J. Bacteriol*. 189: 4465-4472.

Stevenson, C.S., Capper, E.A., Roshak, A.K., Marquez, B., Eichman, C., Jackson, J.R., Mattern, M., Gerwick, W.H., Jacobs, R.S., and Marshall, L.A. (2002a) The identification and characterization of the marine natural product scytonemin as a novel antiproliferative phamacophore. *The Journal of Pharmacology and Experimental Therapeutics*. 303:2, 858-866.

Stevenson, C.S., Capper, E.A., Roshak, A.K., Marquez, B., Grace, K., Gerwick, W.H., Jacobs, R.S., and Marshall, L.A. (2002b) Scytonemin – a marine natural product inhibitor of kinases key in hyperproliferative inflammatory diseases. *Inflammatory Research*. 51, 112-114.

Sun, W., Shahinas, D., Bonvin, J., Hou, W., Kimber, M.S., Turnbull, J., and Christendat, D. (2009) The crystal structure of Aquifex aeolicus prephenate dehydrogenase reveals the mode of tyrosine inhibition. *J. Biol. Chem*. 284: 13223-13232.

Xie, G., Keyhani, N.O., Bonner, C.A., and Jensen, R.A. (2003) Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiol. Mol. Biol. Rev.* 67: 303-342.

CHAPTER FIVE -

PROBING THE ENZYMATIC POTENTIAL OF SCY1263, A HYPOTHETICAL PROTEIN FROM THE SCYTONEMIN BIOSYNTHETIC GENE CLUSTER IN *NOSTOC PUNCTIFORME* ATCC 29133

Abstract

Many organisms biosynthesize UV absorbing metabolites to defend themselves against UV radiation. In some cases, these metabolites require the use of oxidoreductases such as tyrosinases during their biosynthesis. Cyanobacteria are photosynthetic prokaryotes that often live in environments exposed to high levels of UV radiation. In order to combat this radiation, a number of cyanobacteria produce scytonemin, a dimeric pigmented metabolite. The gene cluster for scytonemin biosynthesis in *N. punctiforme* ATCC 29133 contains a hypothetical protein (NpR1263) with sequence similarity to a tyrosinase. This protein was predicted to catalyze the dimerization of two scytonemin monomers to form the final dimeric structure. In this study, the function of the NpR1263 gene product is examined through comparison with a previously described mushroom tyrosinase. These analyses provide the first experimental evidence that Scy1263 can function as an oxygenase enzyme such as a tyrosinase with limited substrate specificity. However, this recombinant protein did not result in the formation of scytonemin when incubated with a synthetic derivative of the scytonemin monomer.

Introduction

The total spectrum of the Sun's intense radiation reaching the Earth's surface can both support life and cause mortality (Caldwell *et al*., 2007). The photosynthetic harvesting of visible wavelengths provides biologically available carbon and oxygenates the atmosphere (Mathews *et al*., 2000). However, the sun also emits higher energy wavelengths known as UV radiation. UV radiation is a source of significant biological damage due to mutagenesis of DNA and proteins and inhibition of biological processes such as photosynthesis, nitrogen fixation, and ATP synthesis (Castenholz and Garcia-Pichel, 2000).

The stratospheric ozone layer blankets the Earth and absorbs a significant amount of the UV radiation from the sun. However, UV-A (315-400 nm) and some UV-B (280-315 nm) wavelengths penetrate the ozone layer and have a biological impact at the surface of the planet (Parson, 2003; Rowland, 2006). Many organisms defend themselves from the harmful effects of UV radiation through the production of UV absorbing metabolites. These metabolites include the mycosporine amino acids found in various organisms, scytonemin found in cyanobacteria, the phenylpropanoids of plants, and the melanins found in animals (Cockell and Knowland, 1999). Although these UV absorbing metabolites vary greatly in range of absorbance and chemical structure, they all use aromatic amino acids as biosynthetic precursors, (Winkel-Shirley, 2001; Soule *et al*., 2007; Korner and Pawelek, 1982; Shick and Dunlap, 2002). Aromatic amino acids absorb UV radiation with the highest energy wavelengths. The coupling of these aromatic amino acids can result in a shift of the

UV absorption to larger wavelengths resulting in the screening of environmentally prevalent UV-A and UV-B radiation (Cockell and Knowland, 1999).

Scytonemin is an example of a UV absorbing metabolite that is produced by cyanobacteria involving the coupling of aromatic amino acid precursors (Proteau *et al*., 1993). The biosynthetic gene cluster for this unique molecule possesses genes involved in aromatic amino acid biosynthesis along with other genes potentially involved in the coupling of a tyrosine derived and tryptophan derived precursor (Balskus and Walsh, 2008; Soule *et al*., 2007). Careful examination of the scytonemin biosynthetic gene cluster across cyanobacterial lineages has also shown some genes within the gene cluster to be species specific (Sorrels *et al*., 2009). One of these genes is NpR1263, a hypothetical protein found only in the scytonemin gene cluster of *Nostoc punctiforme* ATCC 29133. Interestingly, this gene has sequence similarity to a tyrosinase involved in melanin biosynthesis in *Streptomyces* sp. (Sorrels *et al*., 2005).

Tyrosinases are members of a large class of enzymes involved in catalyzing the transfer of electrons from one molecule to another known as oxidoreductases (Dewick, 2002). These enzymes are copper containing oxidoreductases known to have both monophenolase and diphenolase activity (Sánchez-Ferrer *et al*., 1995). The copper mediated mechanism of tyrosinase oxygenase activity is diagrammed in **Figure 5.1**. Tyrosinases are commonly studied for their role in melanin biosynthesis. Melanin is a UV-absorbing metabolite found in many organisms including bacteria, fungi, plants and animals derived from the aromatic amino acid precursor tyrosine (Sánchez-Ferrer *et al*., 1995). Tyrosinase catalyzes the oxidation of tyrosine to yield

3,4-dihydroxy-L-phenylalanine (Dopa). Dopa is then converted to indole-5,6-quinone which is used to produce melanin as outlined in **Figure 5.2** (Korner and Pawelek, 1982).

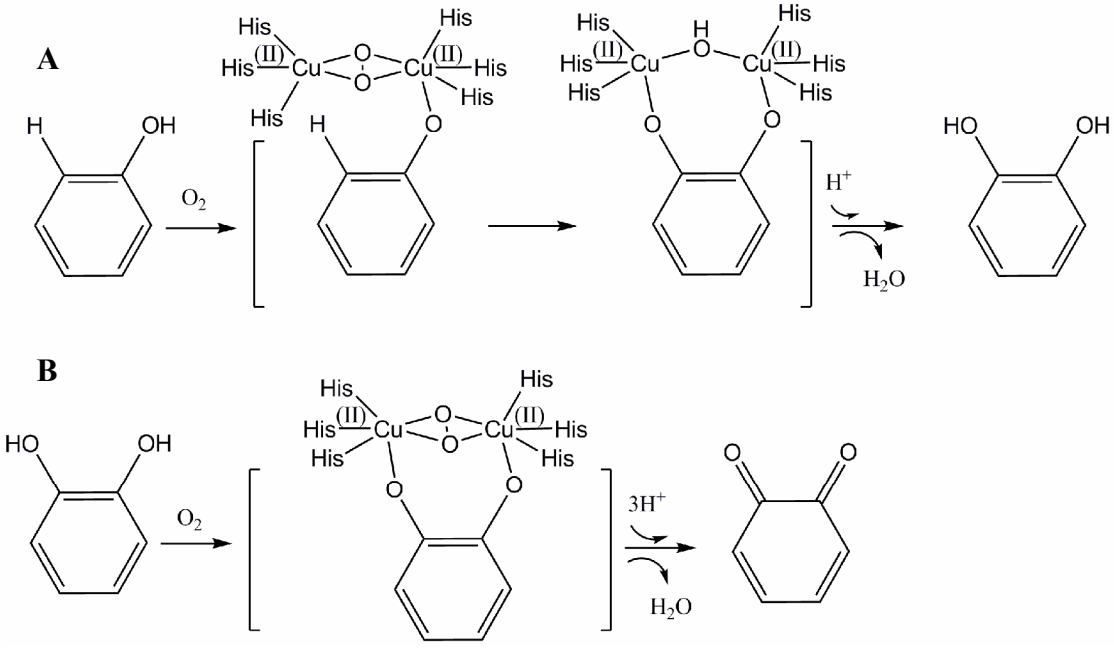


Figure 5.1: Enzymatic mechanism of A) monophenolase and B) diphenolase activity of the tyrosinase enzyme (recreated from Siegbahn, 2003).

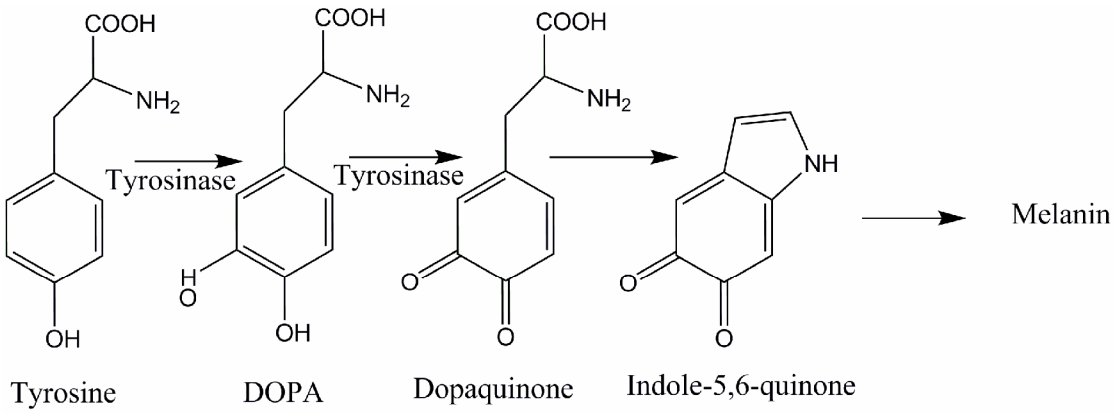


Figure 5.2: Diagram of enzymatic conversion of tyrosine to form melanin (Recreated from Körner and Pawelek, 1982).

Recently, NpR1263, the predicted tyrosinase in *N. punctiforme* ATCC 29133, was speculated to be involved in the proposed coupling of the scytonemin monomers to yield the final dimeric scytonemin structure (**Figure 5.3A**; Balskus and Walsh, 2008). The presence of this gene in only one of the six gene clusters analyzed suggests that its enzymatic function may not be required for the biosynthesis of scytonemin (Sorrels, *et al*., 2009). However, the proposed tyrosinase activity of NpR1263 may still function in the oxygenation of the scytonemin monomer (**Figure 5.3B-C**).

In this study, the function and substrate specificity of a recombinant protein created from NpR1263 is examined in comparison to a commercially available mushroom tyrosinase using a synthetic mimic of the predicted scytonemin monomer.

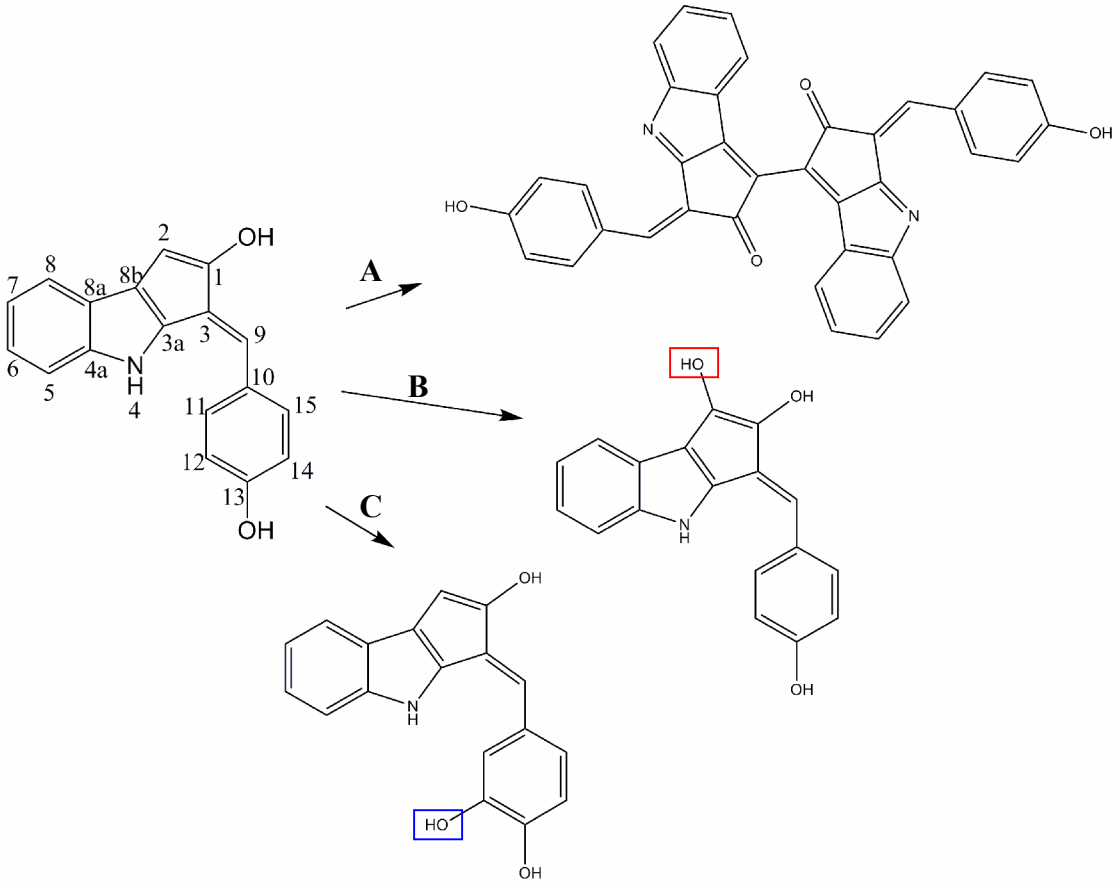


Figure 5.3: Diagram of potential reactions on the predicted scytonemin monomer catalyzed by NpR1263. A) Dimerization to form scytonemin (recreated from Balskus and Walsh, 2008), B) Oxidation at the 2 position outlined in red, C) Oxidation on the phenol outlined in blue. Numbering from Proteau *et al*., 1993.

Materials & Methods

Cyanobacterial Strains and Culture Techniques

The cyanobacterium *Nostoc punctiforme* ATCC 29133 was obtained from the American Type Culture Collection (ATCC). A culture was maintained in unialgal condition in liquid BG-11 freshwater media at 29^{o}C under a light intensity of approximately 19 µmol m^{-2} s^{-1} and a light/dark cycle of 16 h/8 h.

General Experimental

Mushroom tyrosinase (polyphenol oxidase) was obtained from Worthington Biochemical Corp. 4-Hydroxyanisole, L-tyrosine, and 4-hydroxybenzoate were obtained from Sigma. 4-Hydroxyanisole and 4-hydroxybenzoate were resuspended in 50 mM sodium phosphate buffer pH 6.5 to a final concentration of 10 mM. L-tyrosine was resuspended in 50 mM sodium phosphate buffer plus 20% (v/v) N,N'-dimethylforamide (DMF) to a final concentration of 10 mM due to poor solubility. 3-methyl-2-benzothiazolinone hydrazone (MBTH) was obtained from Sigma and resuspended in 50 mM sodium phosphate buffer plus 20% (v/v) DMF to a final concentration of 10 mM also due to poor solubility. The scytonemin monomer derivative (ScyM) and the ortho-dihydroxy scytonemin monomer derivative (ScyM-diOH) were synthesized in the Gerwick laboratory (Suyama, 2009). Both ScyM and ScyM-diOH were resuspended in methanol to a final concentration of 7.3 mM.

Liquid chromatography and mass spectrometry analyses were carried out on a Finnigan LCQ Advantage mass spectrometer. UV/Vis absorption spectra and enzymatic assay results were obtained on a Beckman-Coulter DU800 spectrophotometer.

Cloning, Expression and Purification of Scy1263

Genomic DNA was isolated from *Nostoc punctiforme* ATCC 29133 using DNA Wizard Genomic DNA Purification kit (Promega). Scy1263 was amplified from

genomic DNA using PFX50 Taq polymerase (Invitrogen) according to the manufacturer's recommendations. The following primers were used to amplify a 1,227 bp region of DNA: 5'-GGGGGCCATGGGAAAACTCCTGCTAAAATCTG-3' and 5'- CCCCCCTCGAGTCTTTGCGTTTTTCTTTCCCA -3'. PCR products were cloned in-frame with a N- and C- terminal 6x His fusion and an N-terminal GST fusion at the NcoI and XhoI sites of the PGS21a-6 expression vector (GenScript Corp.). The construct was verified by sequencing (Seqxcel, San Diego CA). Overexpression of protein was carried out in *Escherichia coli* BL21 (DE3). Overnight cultures expressing Scy1263 were diluted 5:100 in 1 L Luria-Bertani (LB) broth containing 100 μg/mL ampicillin and grown at 37^{o}C for 2.5 h prior to isopropyl β-D-1-thiogalactopyranoside (IPTG) induction and then grown 18-20 h at 16^{o}C with shaking (180 rpm). Cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris-HCL pH 8.0, 300 mM NaCl, and 20 mM imidazole), and lysed by sonication for five 10 s bursts at 56 mAmps using a UPC 2000U sonicator (Ultrasonic Power Corporation). Recombinant protein was purified using nickel chelate chromatography (Qiagen). Protein lysates were incubated with Ni-agarose beads for 2 h at 4^{o}C. The protein-Ni-agarose slurry was then washed three times with wash buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, and 50 mM Imidazole), and the protein was eluted with elution buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, and 250 mM Imidazole). Purified proteins were dialysized overnight at 4^{o}C against 50 mM sodium phosphate buffer pH 6.5 using Slide-A-Lyzer Dialysis Cassettes with a molecular weight cutoff of 10,000 (Thermo-Scientific). The GST fusion tag was cleaved from

the recombinant protein using TAG-OFF rEK Cleavage Capture kit (Novagen) according to the manufacturer's recommendations and the resulting cleavage product was used for size comparison purposes with the uncleaved dialysis proteins in SDS-PAGE gel analyses. The dialysized proteins were aliquoted, flash frozen in liquid nitrogen, and stored at -80^{o}C. Protein concentration was determined by Bradford protein assay, and protein purity was assessed by analysis of eluted protein on a 4-20% Tris-HCL SDS-PAGE gel (Biorad) followed by staining with Colloidal Coomassie blue G-250.

UV/Vis Spectrophotometric Assays

Spectrophotometric assays were carried out by measuring the appearance or disappearance of a product at room temperature in the reaction medium using a UV-Vis spectrophotometer. Reference cuvettes containing all the components except the substrate were shown to have no activity (Data Not Shown). Isolation and purification of the empty PGS21a-6 expression vector product expressed in *E. coli* BL21 DE3 and following the same purification procedure described above also resulted in no activity indicating all enzymatic activity is due to the NpR1263 recombinant gene product (Data Not Shown). The initial studies resulting in the identification of the contaminated ScyM were completed as follows: 25 μg of mushroom tyrosinase or 48 μg of Scy1263 eluted protein, 565 μL 50 mM sodium phosphate buffer pH 6.5, 400 μL 10 mM MBTH solution, and 100 nM 4HA or 364 nM ScyM. pH dependence studies of the enzymes were completed as follows: 25 μg of mushroom tyrosinase or 48 μg of Scy1263, 36 μM ScyM, and 950 μL 50 mM sodium phosphate buffer pH 5.3, 5.9, 6.5,

or 7.1. Substrate specificity assays were completed as follows for the substrates L-tyr, 4HA, and 4-HBA with the mushroom tyrosinase enzyme: 25 μg of mushroom tyrosinase, 570 μL 50 mM sodium phosphate buffer pH 5.3, 400 μL 10 mM MBTH solution, and 35 μM substrate. For the ScyM substrate reaction conditions were 25 μg of mushroom tyrosinase, 970 μL 50 mM sodium phosphate buffer pH 5.3, and 35 μM substrate. Substrate specificity assays were completed as follows for the substrates L-tyr, 4HA, and 4-HBA with the Scy1263 enzyme: 43 μg of Scy1263 eluted protein, 545 μL 50 mM sodium phosphate buffer pH 7.1, 400 μL MBTH solution, and 35 μM substrate. For the ScyM substrate reaction conditions were 43 μg of Scy1263 eluted protein, 950 μL 50 mM sodium phosphate buffer pH 7.1, and 35 μM substrate.

Enzymatic Product Assay

Ten cuvettes were used for both Scy1263 and the mushroom tyrosinase enzyme. Each cuvette contained 950 μL 50 mM sodium phosphate buffer (pH 5.3 for mushroom tyrosinase and pH 7.1 for Scy1263), 54 μM ScyM, and either 25 μg of mushroom tyrosinase or 95 μg of Scy1263. This reaction was repeated until sufficient material was obtained. For each set of ten, five contained the same amount of enzyme denatured by boiling for 30 minutes. The reactions were left overnight and then extracted with 1:1 ethyl acetate. The ethyl acetate was evaporated to dryness and resuspended in methanol to a concentration of 1 mg/ml. The resulting mixture was filtered through a 0.2 μm filter and directly injected into the mass spectrometer. The resulting ions were analyzed in negative ionization mode. Liquid chromatography mass spectrometry (LCMS) was completed after solid phase purification using a Strata

C-18E Sep-Pak and eluted with 100% MeOH. The extracts were subjected to C18 RP-HPLC (Phenomenex Synergi 4u Fusion-RP 80 250 x 4.60 mm RP-HPLC column, 4 μm, gradient 70:30 MeOH/H_2O to 100% MeOH over 10 min, 20 min 100% MeOH, 100% MeOH to 70:30 MeOH/H_2O over 15 min; total scan from λ_{max} = 200 - 800 nm with a flow rate of 0.8 mL/min or 0.5 mL/min).

Results & Discussion

Scytonemin's unique dimeric chemical structure has been shown to be derived from the enzymatic coupling of an indolic and phenolic precursor in *N. punctiforme* ATCC 29133 (Balskus and Walsh, 2008). However, no enzymatic function has been shown to be involved in the coupling of the scytonemin monomers to form the final dimeric chemical structure. It has even been suggested that this dimerization may occur non-enzymatically. In an effort to better understand this coupling reaction, the proposed scytonemin monomer (ScyM) was synthesized in the Gerwick laboratory (**Figure 5.4**; Suyama, 2009). The availability of ScyM as an enzymatic substrate combined with the previous speculation that the NpR1263 gene product may be involved in an oxygenase reaction led to an interest in studying the interaction of ScyM with the NpR1263 gene product and other tyrosinase enzymes.

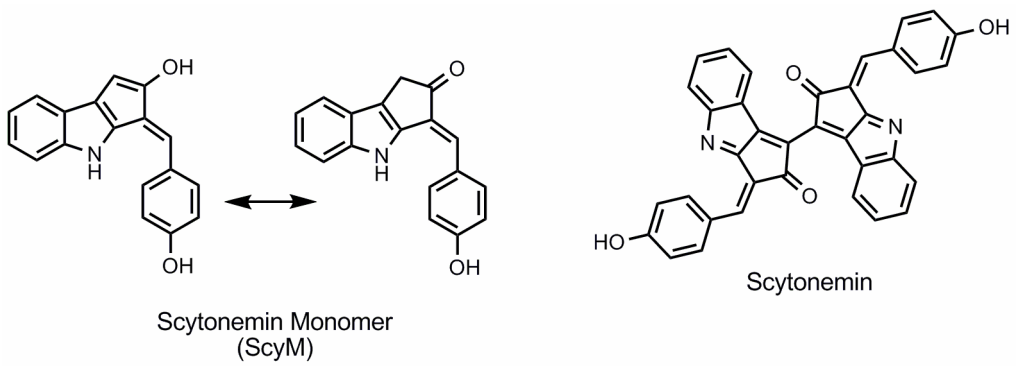


Figure 5.4: Chemical structures of the scytonemin monomer (ScyM) used in this study and scytonemin.

A commercially available mushroom tyrosinase was obtained and shown to have significant activity when incubated with 4-hydroxyanisole (4HA). This activity was monitored using a UV/Vis spectrophotometric assay previously described (Espin *et al*., 2000). This colorimetric assay involves the oxidation of the phenolic substrate resulting in a diphenol. This diphenol is then enzymatically converted to an o-quinone that is the substrate for nucleophilic attack by the amino group of MBTH (**Figure 5.5A**). The resulting product often yields a significant change in absorption, such as the change in color from clear to orange in the case of 4HA, which can be monitored through UV/Vis spectrophotometry at 492 nm. This assay was also used to analyze the formation of product in the reaction between the mushroom tyrosinase and ScyM that was monitored at 535 nm. This reaction resulted in the formation of a purple pigment. The resulting activity of mushroom tyrosinase with 4HA and ScyM using the MBTH coupled spectrophotometric assay is shown in **figure 5.5**.

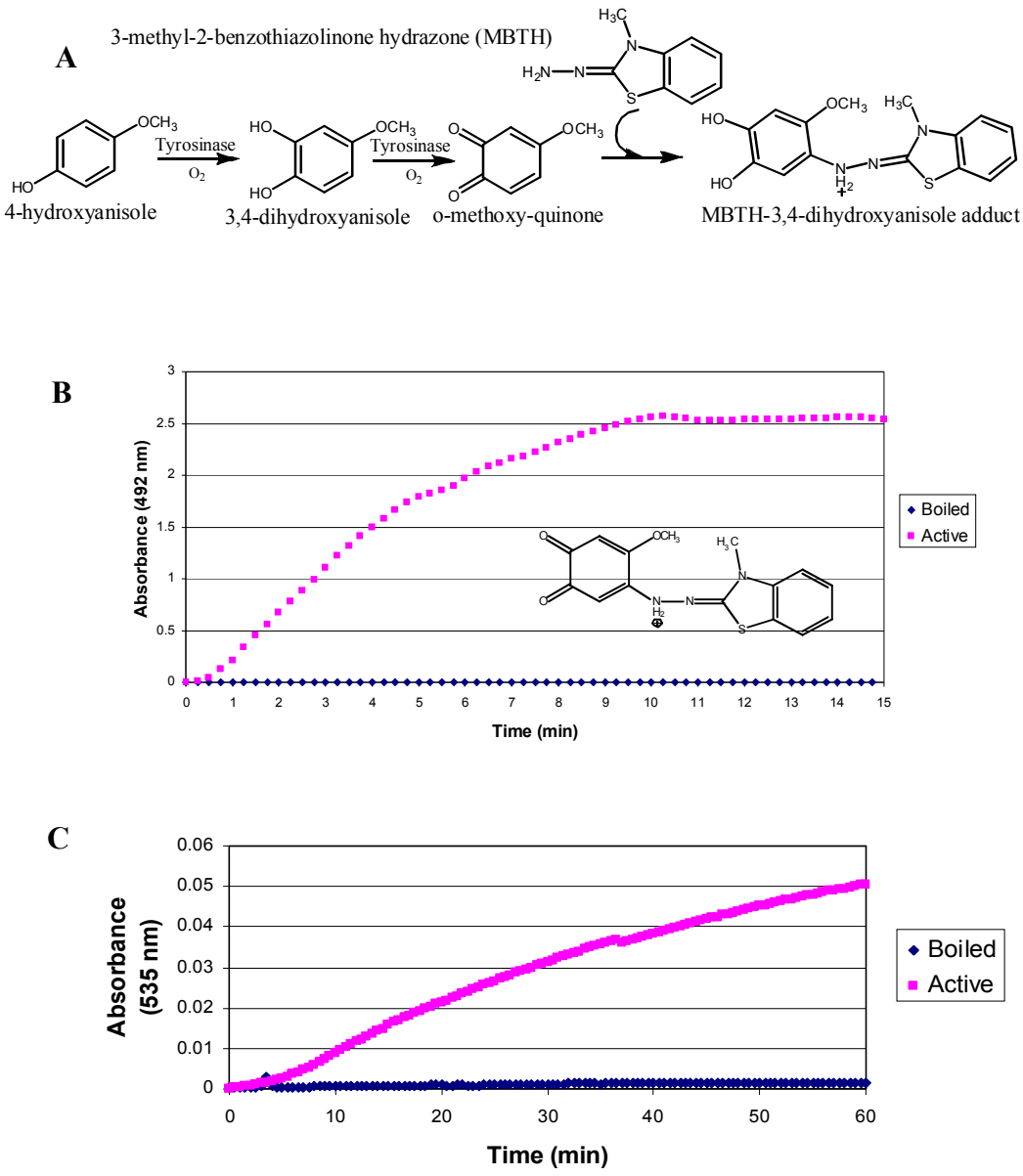


Figure 5.5: MBTH coupled spectrophotometric assay: A) Diagram of mushroom tyrosinase reactivity with 4HA and resulting MBTH coupled product. Results for a boiled and active mushroom tyrosinase enzyme reaction for B) 4HA and C) ScyM (Espín *et al*., 1998).

Although the enzyme appears to react with both substrates, there is a significant difference in the rate of the change in absorbance between these substrates. While 4HA is known to be good substrate for the mushroom tyrosinase, ScyM does not appear to be very reactive with the mushroom tyrosinase (Espín *et al*., 1997). However, even the low activity of the mushroom tyrosinase on ScyM suggested that this tyrosinase was capable of reacting with the predicted scytonemin monomer substrate.

These initial results with the mushroom tyrosinase could subsequently be used as a comparison for the analysis of the function of Scy1263. Scy1263 is the resulting recombinant protein expressed from an expression construct created using the NpR1263 gene from *N. punctiforme* ATCC 29133. This construct yields a purified protein product of between 37 and 50 kDA as expected for the NpR1263 product (45 kDA). This product is the result of non-specific cleaving of the GST-fusion tag during purification as confirmed through the analysis of the cleaved protein product using a GST-fusion tag cleavage kit (**Figure 5.6**).

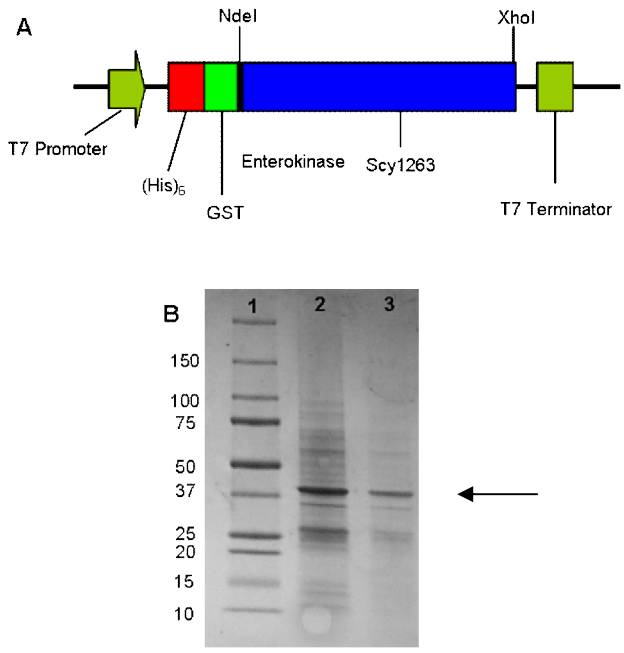


Figure 5.6: Recombinant protein expression of Scy1263. A) Diagram of the construct used in protein expression, B) SDS-PAGE gel stained with colloidal coomassie blue showing the expression of the purified recombinant protein: 1- Precision Plus Dual Color Protein Ladder (Biorad, in kDa), 2- Scy1263 Eluted Protein, 3- Scy1263 with GST Cleaved. Arrow shows protein band representing Scy1263.

The Scy1263 recombinant protein was incubated with ScyM and analyzed using the same MBTH coupled spectrophotometric assay as previously described for the reaction with the mushroom tyrosinase. The reaction of Scy1263 with ScyM also resulted in the formation of a purple pigment. Although both the mushroom tyrosinase and Scy1263 appeared to be reacting with ScyM, analysis of the product by mass spectrometry did not reveal the expected mass for the oxidized ScyM after

reaction with MBTH (*m/z* 524). The viability of the assay was further questioned when a new vial of ScyM was provided and required a different incubation time than previously seen for this substrate to yield the purple pigment.

The synthesis of ScyM required the use 4-hydroxybenzaldehyde (4-HBA) to form the phenolic portion of the molecule (Suyama, 2009). When this synthetic substrate was examined in the MBTH spectrophotometric assay using the mushroom tyrosinase, the product resulted in a purple pigmentation similar to the pigmentation previously seen using the ScyM substrate. These results suggested that although the 4-HBA contaminant was only present in small quantities, the MBTH coupled spectrophotometric assay would not be useful due to the intense color change of this contaminating product.

In order to monitor the enzymatic reaction between the proteins and ScyM, a non-coupled spectrophotometric assay was developed. This assay was designed based on differences between the absorption profile of the boiled enzyme products and the active enzyme products. The intensity of the absorbance at 410 nm was found to decrease in the spectra for the product of the active enzyme reactions (**Figure 5.7**). The contaminant, 4-HBA, did not have an absorption peak at this wavelength; therefore, 410 nm was chosen for monitoring reactivity in the assay.

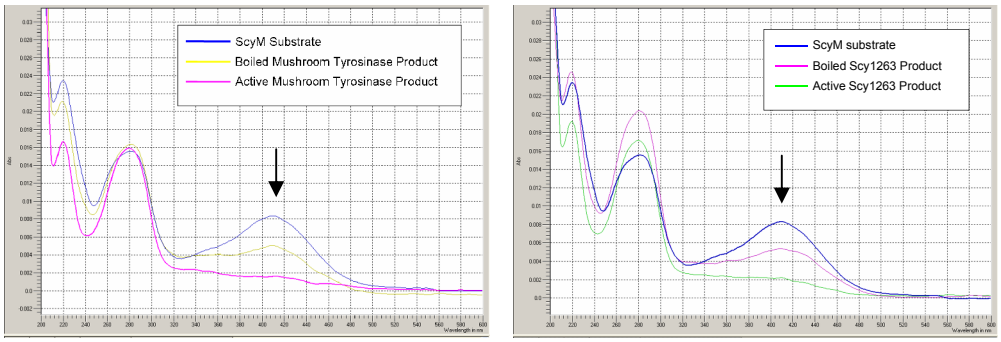


Figure 5.7: UV/Vis absorbance profile for ScyM substrate, boiled protein product and active protein product for A) the mushroom tyrosinase and B) Scy1263. These profiles were used to determine the wavelength for monitoring the reaction of the protein with the ScyM substrate. Wavelength chosen (410 nm) is marked with an arrow.

Once the mushroom tyrosinase and Scy1263 were determined to be reactive with ScyM based on the UV/Vis spectra, the product of these reactions was examined through extraction and analytical analyses. The crude extracts of the enzymatic reactions were analyzed by mass spectrometry through direct injection to reveal a shift in the parent ion of ScyM at *m/z* 274 to a new parent ion at *m/z* 290. This mass shift was seen only in the extracts resulting from reaction with active enzyme (**Figure 5.8**). This shift of 16 mass units suggests that the product of this reaction is ScyM with the addition of an oxygen atom.

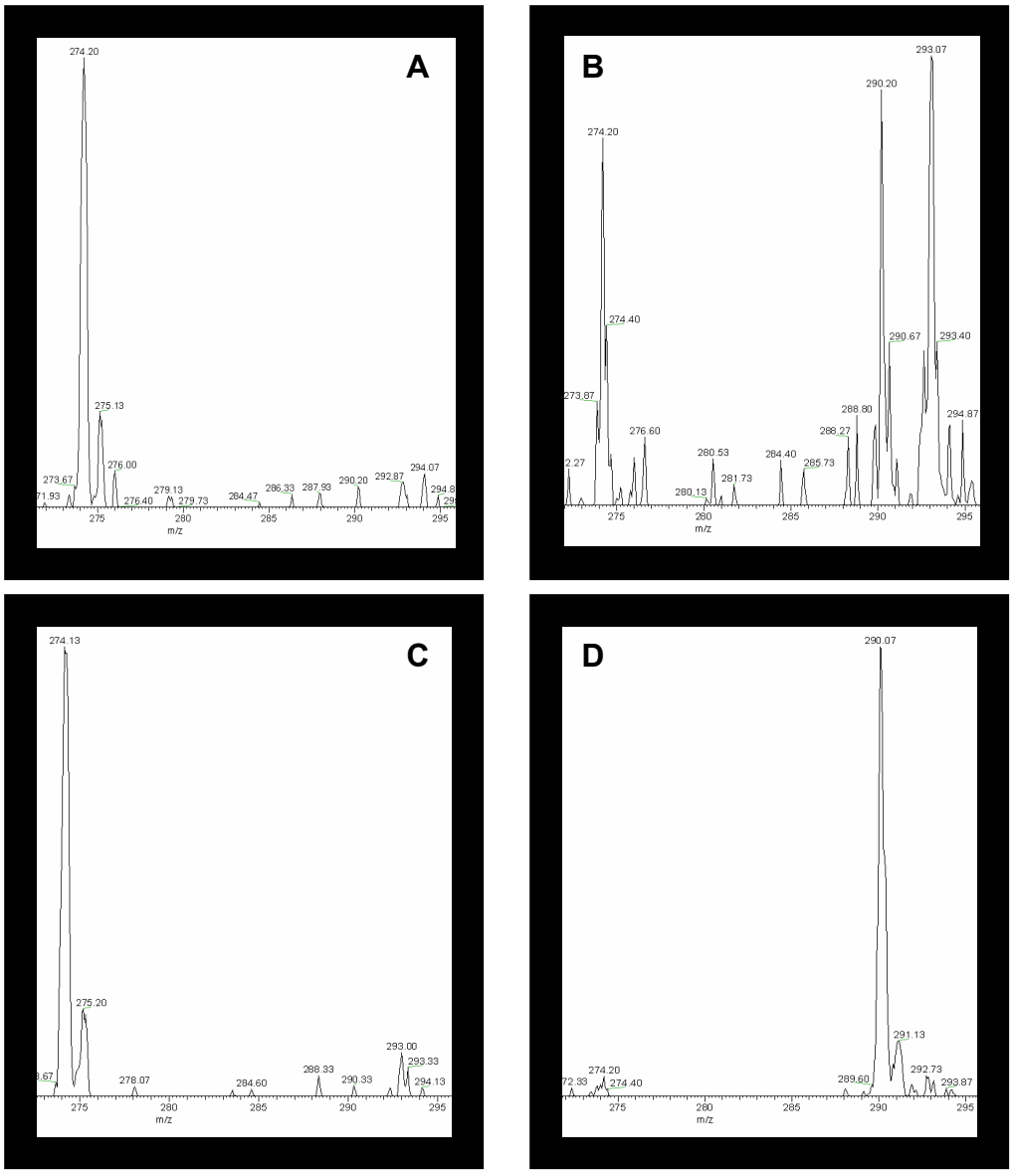


Figure 5.8: Mass spectra of the reaction product for both a boiled and active form of mushroom tyrosinase and Scy1263 when incubated with ScyM and analyzed by mass spectrometry through direct injection. A) Boiled mushroom tyrosinase, B) active mushroom tyrosinase, C) boiled Scy1263, D) active Scy1263.

In order to determine that the molecules isolated from both reactions with *m/z* 290 were the same, these extracts were analyzed using LCMS. The total scan from λ_{max} = 200 to 800 nm revealed three retention times containing the *m/z* 290 for both the mushroom tyrosinase and Scy1263 extracts. The LCMS trace for boiled and active extracts from the mushroom tyrosinase and Scy1263 are shown in **Figure 5.9**. The prominent peak in the spectra at the retention time of about 12.05 min represents the residual ScyM (*m/z* 274). Three other peaks with retention times of about 10.35, 10.64, and 10.95 min were shown to have the mass peak *m/z* 290. Although these three peaks did not appear to be clean products, the pattern of these mass peaks is similar between the active mushroom tyrosinase and Scy1263 extracts. Based on the analysis of the peaks with similar retention times in the boiled control, underlying molecules from a proposed nonenzymatic reaction appear to elute with the *m/z* 290 molecule. These nonenzymatically derived molecules give rise to many of the extra mass peaks including *m/z* 274 and *m/z* 304. The similarities between the LCMS peaks for the active enzymes with retention times 10.35, 10.64, and 10.95 min were also shown based on similar UV/Vis absorbance profiles for each peak (data not shown).

The mass peak *m/z* 290 is only present in the enzymatic reactions containing active enzyme; therefore, this mass is proposed to be the reaction product of the enzymatic reaction for both the mushroom tyrosinase and Scy1263. The presence of a mass peak *m/z* 290 at three different retention times is interesting and may indicate oxidation at various locations on ScyM. The mass spectra and UV/Vis profile

similarities between the peaks at the retention times of approximately 10.35, 10.64, and 10.95 min for the mushroom tyrosinase and Scy1263 indicate that these enzymes are producing similar reaction products.

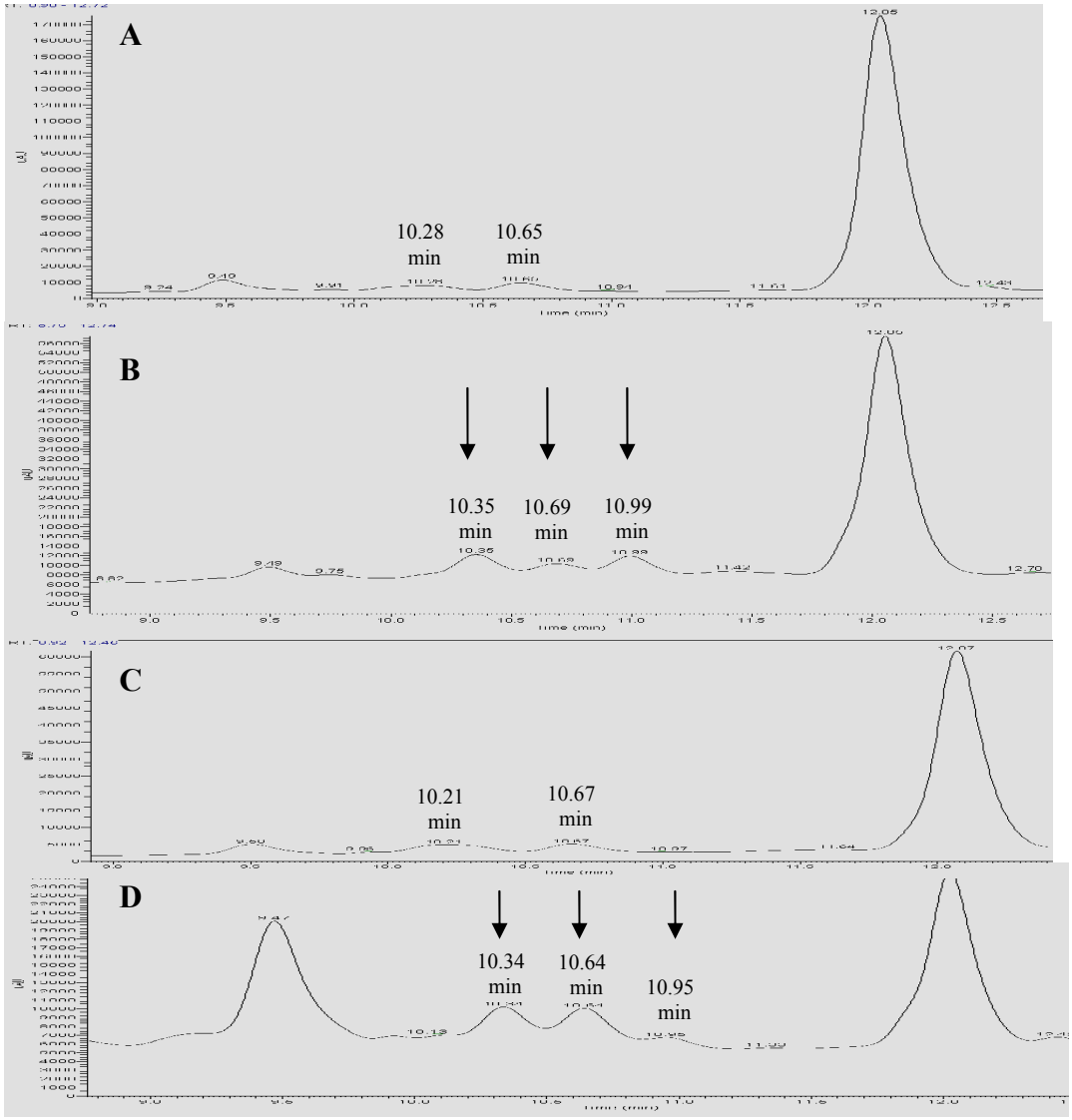


Figure 5.9: Spectral profile of crude extracts obtained from enzymatic reactions of ScyM with mushroom tyrosinase and Scy1263 by LCMS. A) Boiled mushroom tyrosinase, B) active mushroom tyrosinase, C) boiled Scy1263, D) active Scy1263. Retention times found to have *m/z* 290 are marked with black arrows.

The analytical data obtained indicates that the mushroom tyrosinase and Scy1263 are both producing the same product during the enzymatic reaction with *m/z* 290. The mushroom tyrosinase has been characterized as having both monophenolase and diphenolase activity on smaller phenolic molecules (Sánchez-Ferrer *et al*., 1995). In the case of the ScyM substrate, the mushroom tyrosinase is predicted to function as a monophenolase to catalyze the addition of an oxygen to the phenol ring forming a diphenol (**Figure 5.10**).

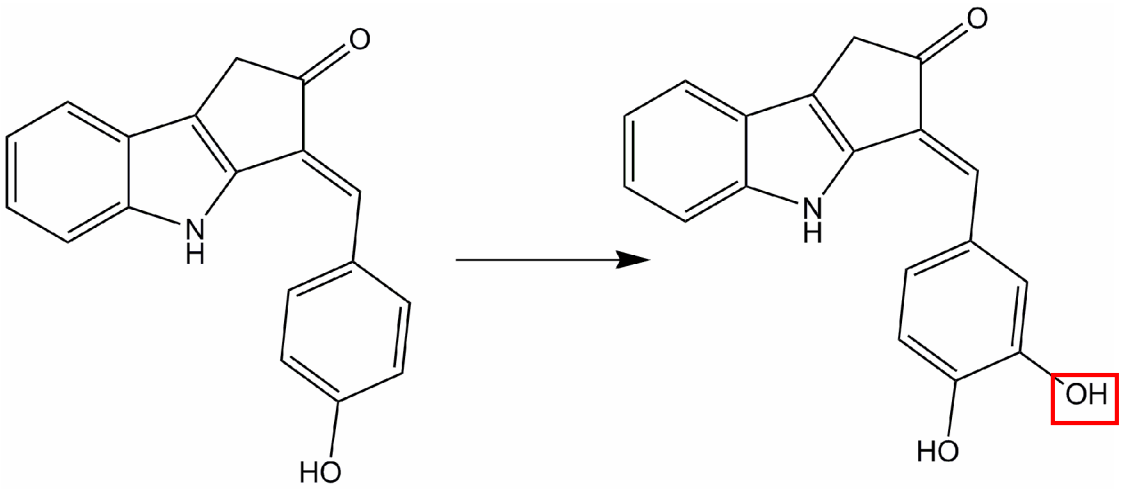


Figure 5.10: Chemical structures of ScyM and the predicted product of the mushroom tyrosinase reaction based on known enzymatic function (Sánchez-Ferrer *et al*., 1995). Red box outlines the addition of a hydroxyl to the phenol ring.

The presence of an additional hydroxyl group on the phenol portion of ScyM is further supported by the mass spectrometry fragmentation data for *m/z* 290 (**Figure 5.11**). The fragmentation pattern shows a number of fragments associated with

common fragment types including M•-18 (*m/z* 272) for H_2O, M•-28 (*m/z* 262) indicative of a loss of C_2H_5 or CO, and M•-42 (*m/z* 248) indicative of a loss of C_3H_6 or C_2H_2O (Crews *et al*., 1998). The fragments with *m/z* 168 and *m/z* 181 are both indicative of the loss of the diphenol portion of ScyM as outlined in **figure 5.11A**. A similar fragmentation is seen from the *m/z* 290 mass peak from the product of the mushroom tyrosinase reaction. These fragments provide support for the presence of the additional hydroxyl group on the phenol portion of ScyM giving rise to *m/z* 290.

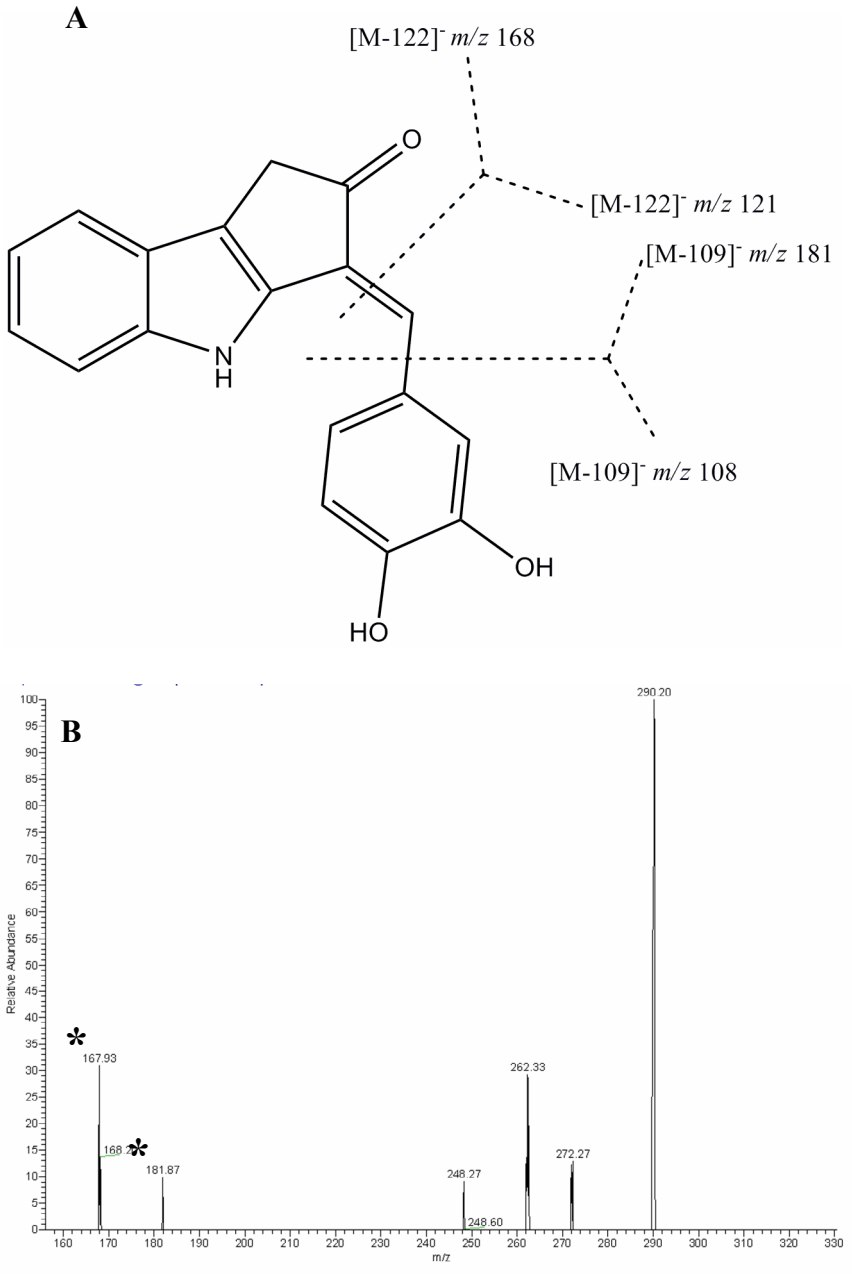
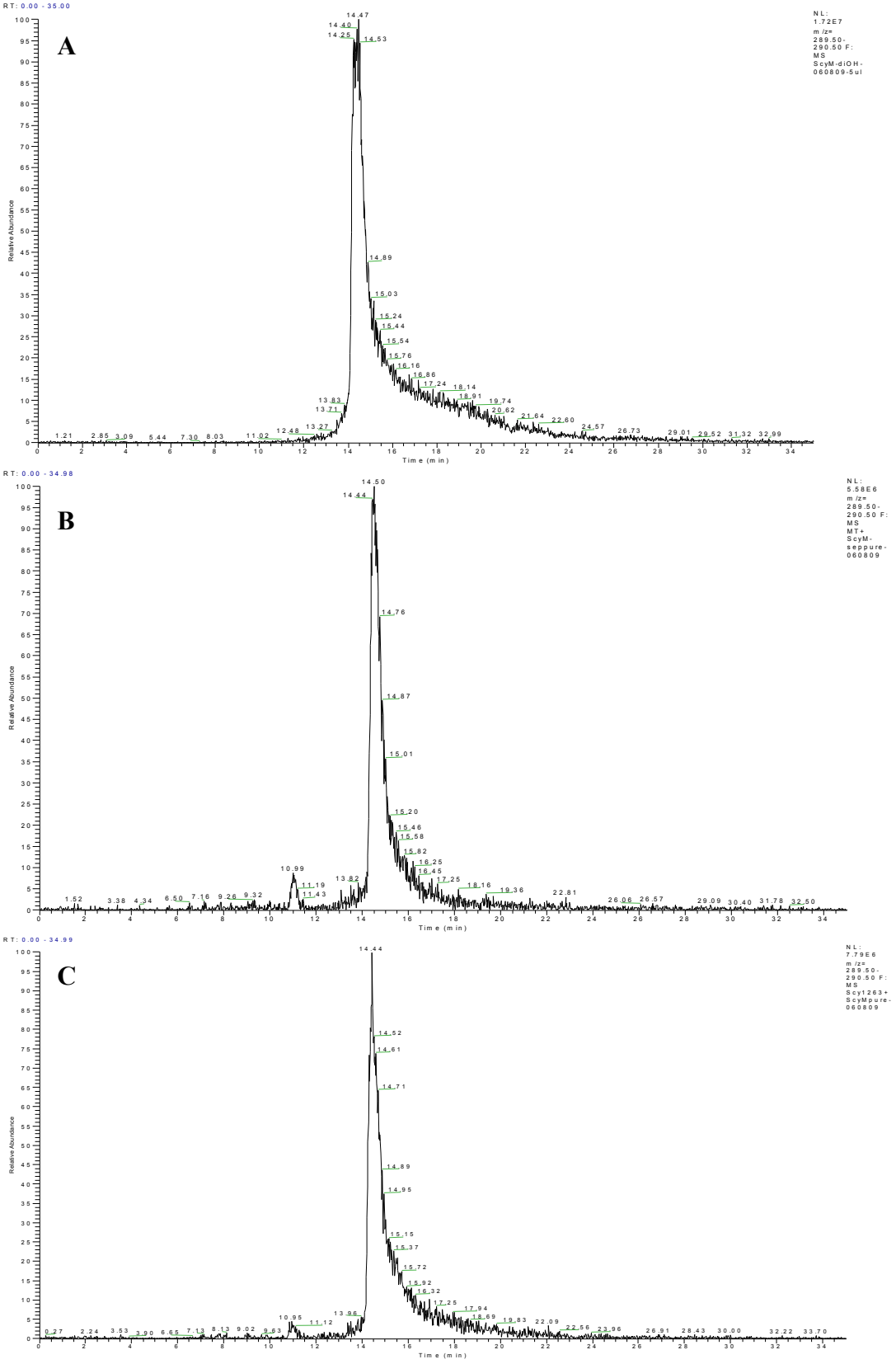


Figure 5.11: Mass spectrometry fragmentation pattern and associated masses for *m/z* 290 obtained from the reaction of Scy1263 with ScyM. A) Chemical structure of the predicted diphenol form of ScyM with important fragmentations marked by dotted lines. B) Fragmentation pattern and associated masses for *m/z* 290 with important fragmentations marked by asterisks. Data obtained in negative ionization mode.

In order to confirm the formation of the o-diphenol form of the scytonemin monomer as predicted from the previous results for both the mushroom tyrosinase and Scy1263, a synthetic standard (ScyM-diOH) was synthesized in the Gerwick laboratory (Suyama, 2009). This standard, the product of the mushroom tyrosinase reaction and the product of the Scy1263 reaction were analyzed by LCMS to reveal similar retention times of about 14.5 minutes (**Figure 5.12**). These results confirm the formation of ScyM-diOH by both the mushroom tyrosinase and the recombinant protein designed from *N. punctiforme* ATCC 29133, Scy1263.

Figure 5.12: Selective ion scan in negative ionization mode for *m/z* 289.5-290.5 used to confirm the formation of ScyM-diOH by mushroom tyrosinase and Scy1263. The A) synthetic ScyM-diOH, B) mushroom tyrosinase product, and C) Scy1263 product all have similar retention times of 14.5 minutes.


A
RT: 0.00 - 35.00
NL: 1.72E7 m/z= 289.50-290.50 F: MS ScyM-diOH-060809-5ul
Relative Abundance
Time (min)
B
RT: 0.00 - 34.98
NL: 5.58E6 m/z= 289.50-290.50 F: MS MT+ ScyM-seppure-060809
Relative Abundance
Time (min)
C
RT: 0.00 - 34.99
NL: 7.79E6 m/z= 289.50-290.50 F: MS Scy1263+ ScyMpure-060809
Relative Abundance
Time (min)

The oxygenase activity of both the mushroom tyrosinase and Scy1263 on ScyM led to an interest in the similarities and differences between these two enzymes. Previous studies have shown that the activity level of tyrosinase enzymes can vary based on the pH of the reaction medium (Espín *et al*., 1998). In order to better understand the effects of pH on the reaction of these enzymes with ScyM, the reaction was studied spectrophotometrically at varying pHs (**Figure 5.13**).

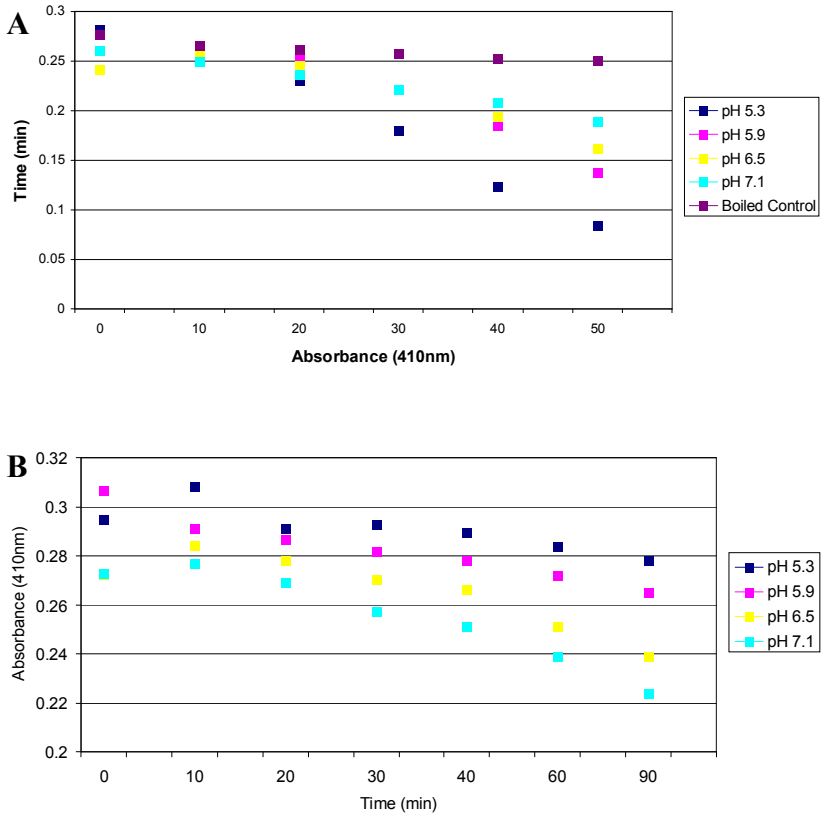


Figure 5.13: Spectrophotometric assay results for A) mushroom tyrosinase and B) Scy1263 at varying pH of the reaction buffer.

Interestingly, these two enzymes appear to have very different optimal pHs. The mushroom tyrosinase is clearly shown to function better at a lower pH of about 5.3. An acidic pH has previously been described to be optimal for this enzyme with other substrates (Espín *et al*., 1998). Scy1263 appears to have optimal function at a pH closer to physiological pH of pH 7.1. However, the results from the boiled enzyme control were emitted from these studies due to these reaction conditions causing ScyM to fall out of solution, leading to a build up a solid yellow material at the bottom of the cuvette. Although the loss of ScyM may play a role in the pH results for Scy1263, the reaction conditions containing an active enzyme never resulted in the appearance of solid material in the cuvette.

The apparent difference in the pH profile of both enzymes may be a result of tolerance to variations in these organisms natural environment. Cultivation of mushrooms at higher pHs results in increased contamination and lower productivity, indicating that mushrooms grow and function best at lower pHs. Thus, it is not unexpected that the mushroom tyrosinase would prefer a more acidic pH (Furlan *et al*., 1997). Although certain cyanobacteria are known to tolerate acidic pH, *N. punctiforme* ATCC 29133 grows optimally at a higher pH between 7 and 8 (Fleming and Castenholz, 2008).

Another question concerning these two enzymes was the difference in substrate specificity. Both enzymes were tested in spectrophotometric assays to analyze their activity with four substrates: 4HA, 4-HBA, L-Tyr, and ScyM. After 20

hours, Scy1263 was shown to have reactivity with only two of the substrates, ScyM and 4-HBA. These results are summarized in **Table 5.1**.

Table 5.1: Summary of substrate specificity results for spectrophotometric assays using Scy1263.

Substrate	λ_{max}	Initial Color	Resulting Color	Change in Absorbance between Boiled and Active Enzyme over 20 h	Activity
4-Hydroxyanisole (4HA)	492[a]	Clear	Clear	0	No
L-Tyrosine	507[a]	Clear	Clear	0	No
4-Hydroxybenzaldehyde (4HBA)	535	Clear	Purple	0.1	Yes
ScyM	410	Bright Yellow	Clear Yellow	0.04[b]	Yes

[a](Espín *et al*., 1998), [b]Also supported by mass spectrometry results

The reactivity of the mushroom tyrosinase with all four substrates allowed for a comparison of the initial velocity of this enzyme with equimolar concentrations of these substrates. These results are shown in **Figure 5.14** and summarized in **Table 5.2**.

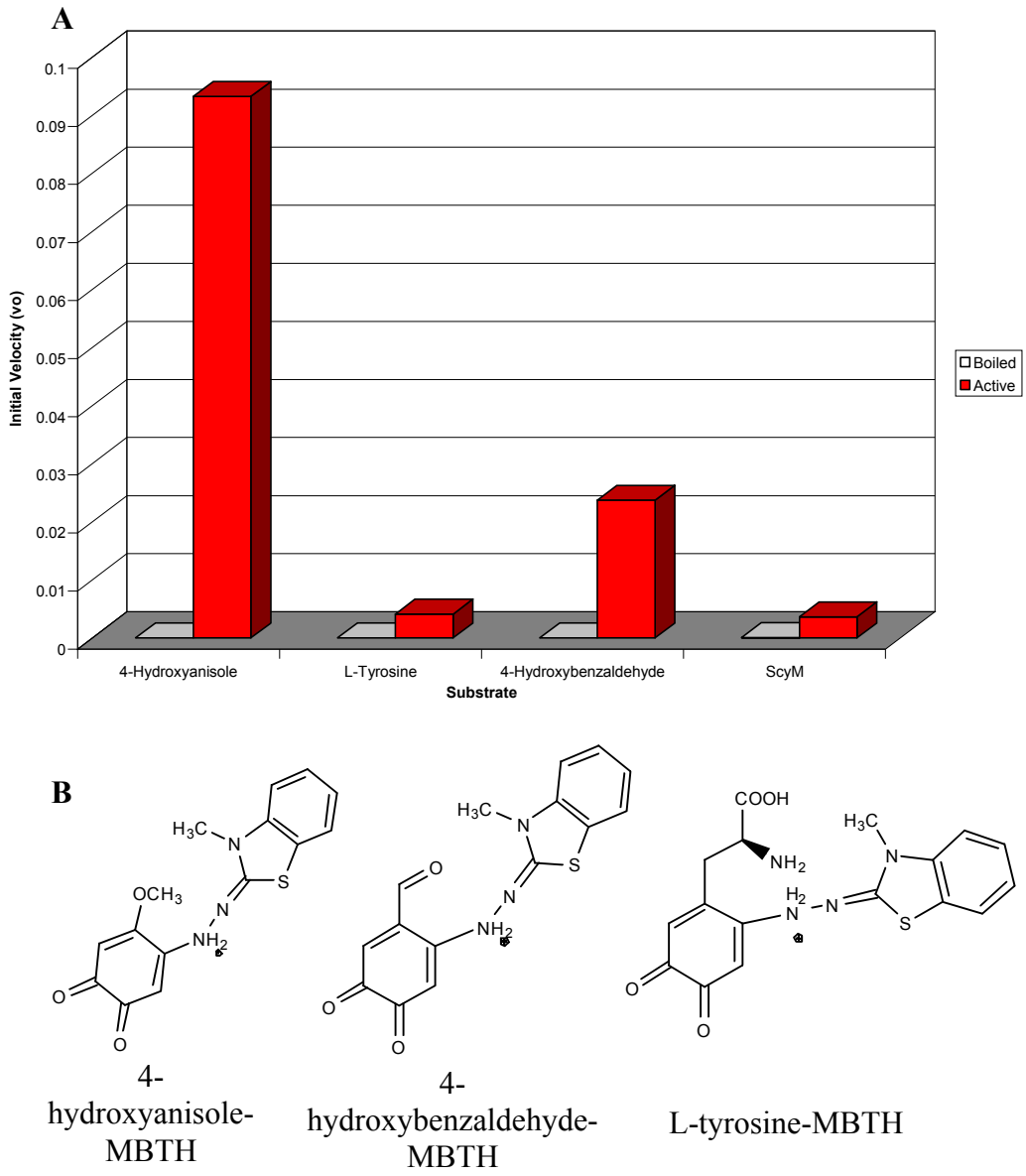


Figure 5.14: Analysis of the substrate specificity for mushroom tyrosinase. A) Bar graph representing the initial velocity for the activity of the mushroom tyrosinase with equimolar concentrations of each of the substrates, 4-hydroxyanisole, L-tyrosine, 4-hydroxybenzaldehyde, and ScyM. B) Chemical structures for predicted reaction products from MBTH coupled reactions. All bars had standard error less than 0.004 for two biological replicates (not shown).

Table 5.2: Summary of spectrophotometric assay results for the mushroom tyrosinase.

Substrate	λmax	Initial Color	Resulting Color	Activity
4-Hydroxyanisole (4HA)	492[a]	Clear	Orange	Yes
L-Tyrosine	507[b]	Clear	Hot Pink	Yes
4-Hydroxybenzaldehyde (4HBA)	535	Clear	Purple	Yes
ScyM	410	Bright Yellow	Clear Yellow	Yes

[a](Espín *et al*., 1998)

The results of substrate specificity for the mushroom tyrosinase clearly show a greater allowance for different substrates; however, the efficiency of the reaction decreases significantly seemingly with the size of the substrate. Both 4HA and 4-HBA appear to be much better substrates based on initial velocity of the reaction at one substrate concentration. The initial velocity for the larger substrates L-Tyr and ScyM is much lower. Interestingly, the same substrate promiscuity is not seen in the reactions with Scy1263. This enzyme appears to be much more specific for ScyM and 4-HBA.

Conclusions

The well studied mushroom tyrosinase is known to accept many phenolic substrates. The oxidation reaction of these phenols is very specific and appears to require the proper spatial docking of the phenols in the active site of the enzyme (Espín *et al*., 2000). In this study, the mushroom tyrosinase is shown to also react with the larger phenolic substrate, ScyM, although with a lower initial velocity compared to other substrates. This study also reveals a newly discovered cyanobacterial tyrosinase, Scy1263, that has a more limited substrate specificity favoring oxidation of larger substrates such as ScyM. Both the mushroom tyrosinase and Scy1263 are shown to react with ScyM to form the ortho-diphenol scytonemin monomer (ScyM-diOH). Scy1263's allowance for fewer substrates is not unexpected as this gene is embedded in a very specific biosynthetic pathway that only seems to be utilized when cyanobacteria are exposed to extremely harsh conditions such as intense UV radiation (Garcia-Pichel and Castenholz, 1991; Sorrels *et al*., 2009).

Although Scy1263 has been proposed to be involved in a coupling reaction to form scytonemin, we found no evidence of this reaction occurring with a similar monomer (Proteau *et al*., 1993; Balskus and Walsh, 2008). This leads to a question of the *in vivo* function of this enzyme in *N. punctiforme* ATCC 29133. Scy1263 may be involved in the formation of a related monomeric species known to occur in other species of *Nostoc* (nostodione A; **Figure 5.15**) or may be involved in the preparation of the monomer for a later coupling reaction. However, the presence of NpR1263 in only one of the genomes shown to contain the scytonemin gene cluster suggests that

this gene may not function in the formation of scytonemin (Sorrels *et al*., 2009). Other genes embedded in the scytonemin gene cluster such as a putative glycosyltransferase (NpR1270) and a series of hypothetical proteins (NpR1271 to NpR1274) including NpR1274 with weak sequence similarity to a gluconolactonase may play a more critical role in preparing the monomer for coupling to form scytonemin. Further studies to analyze the reactivity of Scy1263 with derivatives of ScyM and the reactivity of other gene products associated with the scytonemin gene cluster may lead to insights into the *in vivo* function of this tyrosinase enzyme and the *in vivo* mechanism of scytonemin dimerization.

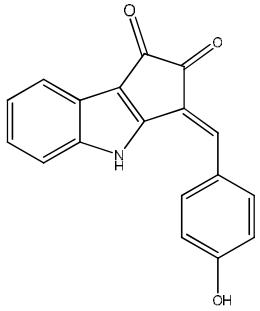


Figure 5.15: Chemical structure of Nostodione A.

Acknowledgements

Chapter 5, in part, is currently being prepared for submission for publication of the material. Sorrels, C.M., Suyama, T., and Gerwick, W.H. The dissertation author was the primary investigator and author of this material.

I would like to acknowledge Tak Suyama for providing the synthetic scytonemin monomer substrate and for being a valuable resource of information during the experimentation. I would like to acknowledge Monica Vodrup for assistance in mathematical calculations. I would also like to acknowledge Dr. Alban Pereirabadilla and Dr. Kevin Tidgewell for additional assistance with NMR at Scripps Institution of Oceanography and the University of California, San Diego.

References

Balskus, E.P. and Walsh, C.T. (2008) Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J. Am. Chem. Soc*. 130: 15260-15261.

Caldwell, M.M., Bornman, J.F., Ballaré, C.L., Flint, S.D., and Kulandaivelu, G. (2007) Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with other climate change factors. *Photochem. Photobiol. Sci*. 6:252-266.

Castenolz, R.W. and Garcia-Pichel, F. (2000) Cyanobacterial responses to UV radiation. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M. Kluwer Academic Publishers: Netherlands, 591- 611.

Cockell, C.S. and Knowland, J. (1999) Ultraviolet radiation screening compounds. *Biol. Rev*. 74: 311-345.

Crews, P., Rodriguez, J., and Jaspars, M. (1998) Optical and chiroptical techniques: ultraviolet spectroscopy. In *Organic Structure Analysis*. Oxford University Press: Oxford, 349-409.

Dewick, P.M. (2002) Secondary metabolites: the building blocks and construction mechanisms. In *Medicinal Natural Products. A Biosynthetic Approach* 2nd ed. John Wiley & Sons, Inc. New York. 7-30.

Espín, J.C., Tudela, J., and García-Cánovas, F. (1998) 4-Hydroxyanisole: the most suitable monophenolic substrate for determining spectrophotometrically the monophenolase activity of polyphenol oxidase from fruits and vegetables. *Anal. Biochem*. 259: 118-126.

Espin, J.C., Varón, R., Fenoll, L.G., Gilabert, M.A., García-Ruíz, P.A., Tudela, J., and García-Cánovas, F. (2000) Kinetic characterization of the substrate specificity and mechanism of mushroom tyrosinase. *Eur. J. Biochem*. 267: 1270-1279.

Fleming, E.D. and Castenholz, R.W. (2008) Effects of nitrogen source on the synthesis of the UV-screening compound, scytonemin, in the cyanobacterium *Nostoc punctiforme* PCC73102. *FEMS Microbiol. Ecol*. 63:301-308.

Furlan, S.A., Virmond, L.J., Miers, D.A., Bonatti, M., Gern, R.M.M., and Jonas, R. (1997) Mushroom strains able to grow at high temperatures and low pH values. *World J. Microbiol. Biotechnol*. 13: 689-692.

Körner, A. and Pawelek, J. (1982) Mammalian tyrosinase catalyzes three reactions in the biosynthesis of melanin. *Science*. 17: 1163-1165.

Mathews, C.K., van Holde, K.E., and Ahern, K.G. (2000) Photosynthesis. *Biochemistry* 3rd edition. Addison Wesley Longman, Inc: USA. 601.

Otávio de Faria, R., Moure, V.R., Lopes de Almeida Amazones, M.A., Krieger, N., and Mitchell, D.A. (2007) The biotechnological potential of mushroom tyrosinases. *Food Technol. Biotechnol*. 45: 287-294.

Parson, E.A. (2003) Early stratospheric science, chlorofluorocarbons, and the emergence of environmental concerns. In *Protecting the Ozone Layer, Science and Strategy*. Oxford University Press. U.S.A. 14-25.

Proteau, P.J., Gerwick, W.H., Garcia-Pichel, F., and Castenholz, R. (1993) The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experentia*. 49, 825-829.

Rowland, F.S. (2006) Stratospheric ozone depletion. *Philos. Trans. R. Soc. B*. 361: 769-790.

Sánchez-Ferrer, Á., Rodríguez-López, J.N., García-Cánovas, F., and García-Carmona, F. (1995) Tyrosinase: a comprehensize review of its mechanism. *Biochemica et Biophysica Acta*. 1247: 1-11.

Shick, J.M. and Dunlap, W.C. (2002) Mycosporine-like amino acids and related gadusols: biosynthesis, accumulation, and UV-protective functions in aquatic organisms. *Annu. Rev. Physiol*. 64: 223-262.

Siegbahn, P.E.M. (2003) The catalytic cycle of tyrosinase: peroxide attack on the phenolate ring followed by O-O bond cleavage. *J. Biol. Inorg. Chem*. 8: 567-576.

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2005) Biosynthesis of scytonemin: a cyanobacterial sunscreen. American Society of Pharmacognosy, Corvallis, OR.

Sorrels, C.M., Proteau, P.J., and Gerwick, W.H. (2009) Organization, Evolution and Expression Analysis of the Biosynthetic Gene Cluster for Scytonemin, A Cyanobacterial Ultraviolet Absorbing Pigment. *Appl. Environ. Microbiol*. 75: 4861-4869.

Soule, T., Stout, V., Swingley, W.D., Meeks, J.C., and Garcia-Pichel, F. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in Nostoc punctiforme ATCC 29133. *J. Bacteriol*. 189: 4465-4472.

Suyama T. (2009) Organic synthesis as an effective approach to chemical, pharmaceutical, and biosynthetic investigations of natural products. Ph.D. Dissertation. University of California, San Diego.

Winkel-Shirley, B. (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol*. 126: 485-493.

CHAPTER SIX -

CONCLUSIONS

The stratospheric ozone layer has been depleted by 3% per year over the past thirty years (McKenzie *et al*., 2003). The decrease in this vital protective atmospheric layer has undoubtedly led to decreases in primary productivity of plants and increases in human health concerns including skin cancer (Gour *et al*., 1997; Henriksen *et al*., 1990). Decreasing global emissions of chlorofluorocarbon compounds is expected to lead to the recovery of the ozone layer in the next 50 years. However, the level of total ozone recovery is predicted to be negatively impacted by the effects of global warming (Rowland, 2006).

Cyanobacteria are photosynthetic gram-negative prokaryotes thought to be among the most ancient organisms on the planet (Cohen and Gurevitz, 2006). Their ability to photosynthesize has long been speculated to have played a role in the oxygenation of the atmosphere, thus allowing for the development of higher life forms (Wynn-Williams *et al*., 2002). The ability of cyanobacteria to survive many environmental regimes over the past 2.8 billion years is a testament to their capacity to adapt to a changing environment (Des Marais, 2000; Paul, 2008). One of these adaptations is the biosynthesis of UV absorbing molecules that function as a protective screen against harmful levels of UV radiation (Cockell and Knowland, 1999).

Scytonemin is an example of a UV absorbing molecule found in the sheaths of cyanobacteria that can prevent 95% of incident UV-A photons from reaching the

subtending cells. This metabolite is a yellow-green pigment described in the sheaths of over 300 species of cyanobacteria (Garcia-Pichel and Castenholz, 1991). Scytonemin's unique dimeric indolic-phenolic chemical structure has led to an interest into its biosynthesis (Proteau *et al*., 1993; Soule *et al*., 2007). An understanding of the basic biology involved in scytonemin's biosynthesis and insights into the unique mechanisms used by cyanobacteria as a defense from UV radiation may lead to exciting new developments with potential applications in biotechnology.

Chapter two uses the powerful tools of bioinformatics and genomics to identify candidate genes involved in scytonemin biosynthesis. This chapter also explores an increase in the level of transcription of these candidate genes based on semi-quantitative RT-PCR analyses. The candidate genes identified in chapter 2 were later shown by another laboratory to be directly involved in the biosynthesis of scytonemin through transposon mutagenesis of *Nostoc punctiforme* ATCC 29133 (Soule *et al*., 2007).

Chapter three provides the first experimental evidence that the genes involved in scytonemin biosynthesis function as a gene cluster. This analysis uses semi-quantitative RT-PCR to show the increased expression of every gene in the cluster after exposure to UV radiation. The chapter also compares the scytonemin gene cluster across six cyanobacterial lineages to reveal the unique components of the cluster between organisms and to suggest the ancient evolutionary origin of scytonemin biosynthesis. These differential transcription studies and genetic analyses reveal a potential regulatory element involved in scytonemin biosynthesis. The

predicted two-component histidine kinase and response regulator found at the end of each cluster may have a specific photoreceptive response to UV radiation. The discovery and understanding of this type of regulatory system could lead to the development of a novel inducible system, important to the advancement of heterologous expression in biotechnology.

An understanding of the scytonemin biosynthetic gene cluster provides access to the enzymatic mechanisms involved in the production of this molecule. One of the genes found in this cluster (Np1276) was shown to catalyze the *in vitro* condensation of indolic and phenolic precursors for scytonemin biosynthesis (Balskus and Walsh, 2008). Chapter four uses a small scale mass spectrometry technique, Matrix Assisted Laser Desorption-Time of Flight (MALDI-TOF) to identify tyrosine and tryptophan as the *in vivo* precursors to scytonemin biosynthesis. Based on the level of isotope incorporation, these precursors are suggested to have a regulatory effect on the endogenous biosynthesis of the amino acid derived precursors. The MALDI-TOF technique also provided a unique near-real-time insight into the induction and rate of the biosynthesis of scytonemin. MALDI-TOF is a powerful new tool that will have a great role in advancing the understanding of cyanobacterial natural products biosynthesis.

The genetic analysis of the scytonemin biosynthetic gene cluster across the cyanobacterial lineages also identified enzymatic features unique to each of the clusters. Chapter five examines the function of one of these unique genes (Np1263) in the *N. punctiforme* ATCC 29133 gene cluster. Using bioinformatics, this gene was

predicted to function as a tyrosinase. A recombinant protein produced using the Np1263 gene was subsequently found to have a similar activity to the well characterized mushroom tyrosinase. The oxygenase activity of these enzymes is shown to react with a synthetic mimic of the predicted scytonemin monomer. This study provides the first *in vitro* result that the scytonemin gene cluster in *N. punctiforme* ATCC 29133 encodes a functional oxygenase protein similar to a tyrosinase. Tyrosinases are important due to their biotechnological uses in nutrition and bioremediation (Otávio de Faria *et al*., 2007). The discovery of a cyanobacterial tyrosinase may lead to advancements in these important areas of biotechnology.

Overall, these studies provide a better understanding of the unusual biosynthesis of scytonemin. Future scytonemin related research will likely reveal its unusual UV specific photoreceptor that is important in transcriptional regulation, the unique enzymatic chemistry involved in its elusive dimerization, and novel scytonemin analogs produced through enzymatic variations in the scytonemin gene cluster across cyanobacterial lineages. Harnessing the many facets of scytonemin biosynthesis may result in the discovery of enzymes and signaling elements with vast biotechnological potential and the engineering of scytonemin biosynthesis to reveal novel molecules with exciting pharmaceutical value. Importantly, developing a better understanding of cyanobacterial UV absorbing metabolites such as scytonemin also provides an important framework for the understanding of potential protective mechanisms against increasing UV radiation at the surface of the planet.

References

Balskus, E.P. and Walsh, C.T. (2008) Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J. Am. Chem. Soc*. 130: 15260-15261.

Cohen, Y. and Gurevitz, M. (2006) The Cyanobacteria – Ecology, physiology and molecular genetics. *Prokaryotes*. 4, 1074-1098.

Des Marais, D.J. (2000) When did photosynthesis emerge on Earth? *Science*. 289: 1703-1705.

Garcia-Pichel, F. and Castenholz, R.W. (1991) Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycology*. 27, 395-409.

Gour, R.K., Pandey, P.K., and Bisen, P.S. (1997) Differential response in damage and repair of wild-type Anacystis nidulans and its UV-B plus heat shock tolerant (UV-HS t) strain under UV-B and heat shock stress. *J. Photochem. Photobiol. B*. 40, 61-67.

Henriksen, T., Dahlback, A., Larsen, S.H.H., and Moan, J. (1990) Ultraviolet-radiation and skin cancer. Effect of an ozone layer depletion. *Photochem. Photobiol*. 51: 579-582.

McKenzie, R.L., Björn, L.O., Bais, A., and Ilyasd, M. (2003) Changes in biologically active ultraviolet radiation reaching the Earth's surface. *Photochem. Photobiol. Sci*. 2: 5-15.

Otávio de Faria, R., Moure, V.R., Lopes de Almeida Amazones, M.A., Krieger, N., and Mitchell, D.A. (2007) The biotechnological potential of mushroom tyrosinases. *Food Technol. Biotechnol*. 45: 287-294.

Paul, V.J. (2008) Global warming and cyanobacterial harmful algal blooms. In *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*. Springer. New York. 619: 239-257.

Rowland, F.S. (2006) Stratospheric ozone depletion. Phil. Trans. R. Soc. B. 361: 769-790.

Soule, T., Stout, V., Swingley, W.D., Meeks, J.C., and Garcia-Pichel, F. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in Nostoc punctiforme ATCC29133. *J. Bacteriol.* 189: 4465-4472.

Wynn-Williams, D.D., Edwards, H.G.M., Newton, E.M., and Holder, J.M. (2002) *Internat. J. Astrobiol.* 1, 39-49.

APPENDIX -

DIFFERENTIAL PIGMENTATION OF *GLOEOCAPSA* SP. UPON EXPOSURE TO UV-A RADIATION

Abstract

Cyanobacteria are gram-negative prokaryotes capable of oxygenic photosynthesis. These organisms produce a wide array of photosynthetic pigments to assist in harvesting sunlight. Cyanobacteria also produce carotenoids in response to oxidative stress induced by their exposure to UV radiation. In this study, the upregulation of three carotenoid pigments in *Gloeocapsa* sp. upon exposure to UV-A radiation is described.

Introduction

The history of the planet is marked with significant abiotic and biotic developments that shaped its evolution (Benton, 2009). One of the most critical biotic developments was oxygenic photosynthesis, a process ultimately resulting in the oxygenation of the planet (Dismukes *et al*., 2001). Oxygenic photosynthesis is the principal way for biologically available carbon to enter the biosphere, a major source of oxygen for the Earth's atmosphere, and the most important source of energy for the planet. Photosynthesis is a specialized biological process that uses the energy from sunlight to convert CO_2 and H_2O into biologically available forms of carbon. This

process also results in the release of oxygen as a byproduct (Mathews *et al*., 2000). Photosynthesis is carried out by plants, algae and cyanobacteria in both terrestrial and oceanic environments. Of the estimated 104.9 gigatonnes of carbon created per year through photosynthesis on Earth, 53.8% is derived from terrestrial sources and 46.2% is derived from oceanic sources (Field *et al*., 1998).

Cyanobacteria, or "blue-green algae," are gram-negative prokaryotes and contribute to the primary productivity of the planet. These organisms harvest light energy from the sun using chlorophyll, and they produce a variety of photosynthetic accessory pigments called phycobiliproteins allowing for an increased spectrum of absorption. Two of the main phycobiliproteins found in cyanobacteria are phycoerythrin and phycocyanin (Glazer, 1977).

The photoautotrophic lifestyle also presents cyanobacteria with an environmental challenge due to the association of visible light and ultraviolet (UV) radiation in the total spectrum of sunlight penetrating the atmosphere. Although the Earth's stratospheric ozone layer screens the most harmful of these UV wavelengths, cyanobacteria are still exposed to significant levels of UV-B (380-315 nm) and UV-A (315-400 nm) radiation (Björn, 2007). UV radiation is responsible for many types of biological damage including mutagenesis of DNA, production of radical oxygen species and inhibition of many physiological processes such as photosynthesis, nitrogen fixation and ATP synthesis (Castenholz and Garcia-Pichel, 2000).

Cyanobacteria often combat these damaging UV wavelengths by localizing UV absorbing molecules in their sheaths to shield the subtending cells. The most

common types of these UV absorbing molecules are scytonemin and the mycosporine-like amino acids (Ehling-Schulz and Scherer, 1999). Cyanobacteria are also known to produce carotenoids in response to UV radiation. These metabolites make up the largest class of pigments and are responsible for a majority of the bright red, yellow, and orange colors seen in pigmented organisms (Hirschberg and Chamovitz, 1994). Carotenoids are well known for their role as antioxidants resulting in the removal of singlet oxygen, triplet chlorophyll and the inhibition of lipid peroxidation (Ehling-Schulz and Scherer, 1999). Cyanobacteria use these antioxidants to limit oxidative stress during periods of exposure to UV radiation (He and Häder, 2002).

The genus *Gloeocapsa* is an example of a cyanobacterium that often lives in environments exposed to extremely harsh conditions including limestone rocks, white roof-tiles, and whitewashed walls. These environments can lead to dessication as well as high levels of visible and UV radiation. *Gloeocapsa* are typically unicellular, with spherical or oblong cells that divide in three planes. The cells are often enclosed in a mucilage capsule that can become pigmented upon exposure to sunlight (Lewin, 2006). In this study, we examine the changes in pigmentation in *Gloeocapsa* sp., a cyanobacterium collected in San Diego, CA, in response to UV-A radiation.

Materials & Methods

Cyanobacterial Strain and Culture Techniques

The cyanobacterium *Gloeocapsa sp.* was obtained and identified by Dr. Ralph Lewin from a sidewalk in San Diego, CA. A culture was maintained in unialgal condition in liquid BG-11 freshwater media at 29^{o}C under a light intensity of approximately 19 μmol m^{-2} s^{-1} and a light/dark cycle of 16 h/8 h. *Gloeocapsa* sp. was grown for approximately 110 days in 2L Fernbach flasks. These cultures were then transferred to large pans for exposure to UV radiation. One pan was exposed to 0.64 mW/cm^2 UV-A radiation (λ_{max} = 365nm) for eight days, while the other control pan remained under normal visible light.

General Experimental

HPLC purification was performed using Waters 515 pumps and a Waters 996 photodiode array detector under computer control using Empower software. LCMS analyses were carried out on a Finnigan LCQ Advantage mass spectrometer attached to a Finnigan Surveyor HPLC system.

Extraction and Isolation

Gloeocapsa sp. was removed from pans by centrifugation resulting in two cyanobacterial pellets: a +UV pellet (52.2 g wet wt.) and a –UV pellet (9.2 g wet wt). Each pellet was exhaustively extracted with CH_2Cl_2/MeOH (2:1) until most of the color was removed. The resulting extracts were concentrated to near dryness and labeled Gloeo-0311-UV and Gloeo-0311+UV. This material was then subjected to solid phase purification using a Strata C-18E Sep-Pak and eluted with 50:50

MeOH/H_2O, 75:50 MeOH/H_2O, 100% MeOH. Only the 100% MeOH fraction contained pigmentation; therefore, the other two fractions were not used in the remainder of the study. Both the Gloeo-0311-UV and the Gloeo-0311+UV fractions were subjected to C18 RP-HPLC (Phenomenex Synergi 4u Fusion-RP 80 250 x 10 mm RP-HPLC column, 4 μm, gradient 70:30 MeOH/H_2O to 100% MeOH over 10 min, 20 min 100% MeOH, 100% MeOH to 70:30 MeOH/H_2O over 10 min; photodiode array detection; λ_{max} = 300, 400, 500 nm). Three compounds were isolated, described as:

Gloeo-0311+UV-A: red pigment; UV (MeOH) λ_{max} = 471 nm; LRESI(-)MS *m/z* [M-H] 572, 586.

Gloeo-0311+UV-B: yellow pigment; UV (MeOH) λ_{max} = 449, 475 nm; LRESI(-)MS *m/*z [M-H] 588, 931, 1544.

Gloeo-0311+UV-C: red pigment; UV (MeOH) λ_{max} = 470, 493 nm; LRESI (-)MS *m/z* [M-H] 586.

Results & Discussion

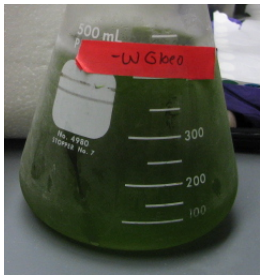
The unicellular cyanobacterial genus *Gloeocapsa* is known to grow in areas prone to high levels of light and extreme dessication, including roof tops, sidewalks and white-washed walls lending them the nickname “Black Algae”(Lewin, 2006). The ability of this genus to survive in such harsh light conditions suggests the use of UV screening molecules for protection against UV radiation. A sample of *Gloeocapsa* sp. obtained from a concrete sidewalk was initially cultivated by Dr. Ralph Lewin.

This strain was found to secrete a red pigment around the edges of the colonies when grown on a windowsill, which was hypothesized to be a response to the high levels of light reaching the cells during cultivation (Lewin, personal communication).

Previous studies using *Gloeocapsa alpicola* have shown that strains containing a greater content of carotenoids are more resistant to higher levels of light, including UV-A radiation. The strains considered wild-type had a lower content of carotenoids and resulted in death upon exposure to UV-A radiation (Buckley and Houghton, 1976). The ability of the *Gloeocapsa* sp. cultures collected by Dr. Lewin to survive under high light intensities suggested that these cultures might also have a resistance to UV radiation due to the presence of carotenoids.

Interestingly, when *Gloeocapsa* sp. was exposed to UV-A radiation for an extended period of time, changes in pigmentation were clearly evident. The exposed cyanobacterial cells were found to be a red-brown color, while the cells maintained under only visible light remained green. Photographs of the crude extracts from both the –UV and the +UV cells clearly show this pigmentation difference (**Figure A.1**).

A



B

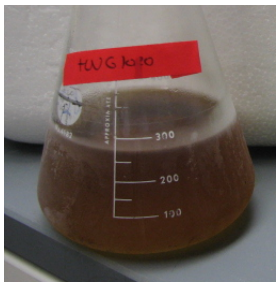


Figure A.1: Photographs of crude extracts obtained from A) –UV control cells and B) +UV cells of *Gloeocapsa* sp.

The obvious difference in the color of the two extracts suggested that *Gloeocapsa* sp. was altering the content of pigmented metabolites in response to UV radiation. The reddish-brown pigmentation in the extract indicated that carotenoids may play a role in the UV response of this species. These extracts were analyzed using HPLC to reveal three significantly different peaks in the +UV extract. These peaks are labeled Gloeo-0311+UV-1, Gloeo-0311+UV-2, Gloeo-0311+UV-3 (**Figure A.2**).

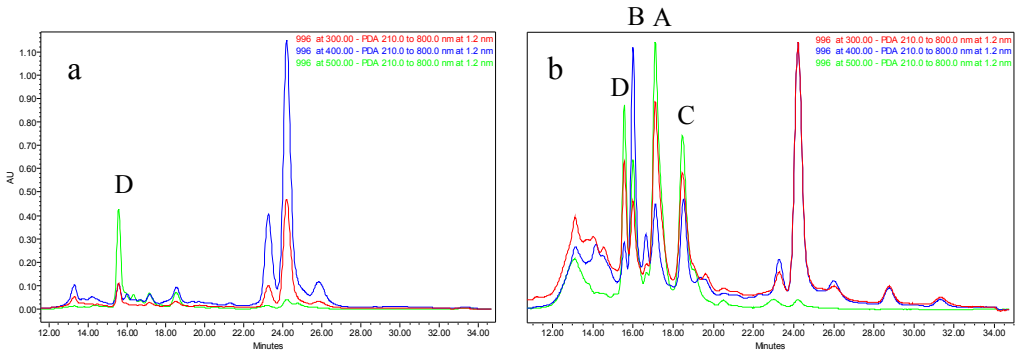


Figure A.2: HPLC traces of extracts from *Gloeocapsa* sp. a) without UV-A radiation exposure, b) with UV-A radiation exposure. Peaks apparent only in the UV-A induced extracts are numbered A, B, and C. Peak D has the characteristic UV/Vis profile of a carotenoid. Detection was measured at λ_{max} = 300, 400, and 500 nm.

All three of the peaks labeled in **figure A.2** were collected to yield three brightly colored pigments with masses between 500-600 *m/z*. The bright pigmentation, experimental masses and the characteristic UV absorption profiles for these pigments indicate that they are all carotenoids. **Figure A.3** diagrams the UV/Vis absorption profile for each of these molecules.

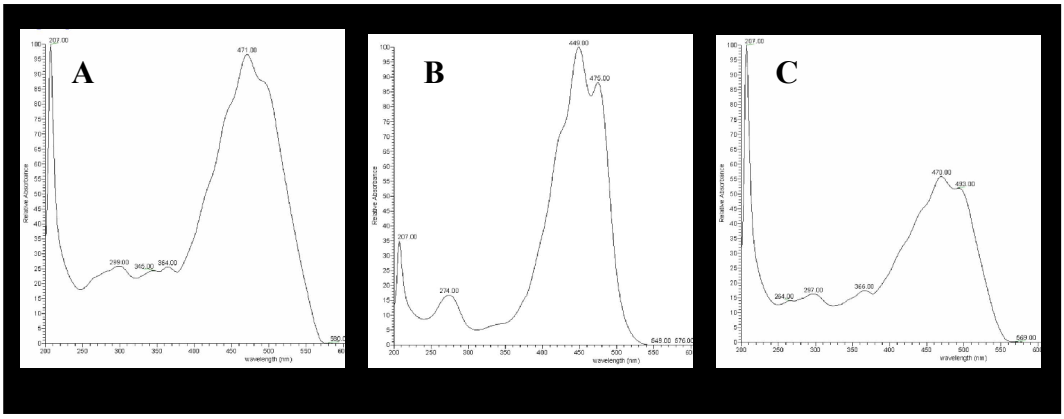


Figure A.3: Absorption profiles for pigments extracted from *Gloeocapsa* sp. after exposure to UV radiation. A) Gloeo-0311+UV-1, B) Gloeo-0311+UV-2, C) Gloeo-0311+UV-3.

Both the position of the absorption maxima and the shape of the profiles are very characteristic of carotenoids (Britton, 1995). Most carotenoids absorb light between 400 and 500 nm as is seen for these three molecules. The position of the absorption maxima and the shape of carotenoid profiles are typically dependent on the length of the chromophore, the geometry of the isomers, and the addition of β-rings. The shape of the profile can also be significantly altered by the presence of a conjugated ketone within the β-ring (Britton, 1995). The UV/Vis spectra for these three pigments have similar shapes; however, the absorption profile of Gloeo-

0311+UV-1 and Gloeo-0311+UV-2 appear to be more similar to each other than either is to Gloeo-0311+UV-3. This suggests that Gloeo-0311+UV-1 and Gloeo-0311+UV-2 have more similar chemical structures.

Two of the three differentially expressed pigments in the *Gloeocapsa* sp. culture exposed to UV radiation were a red to red-orange color. Based on these results, the pigmentation previously seen surrounding the colonies grown on agar plates were probably carotenoids. Future studies using NMR and high resolution mass spectrometry will reveal the identity of these induced carotenoids.

Previous studies have shown that the composition of carotenoids in cyanobacteria can vary from one environmental location to another in response to changing light conditions (Lakatos *et al*., 2001). In this study, I found that the composition of carotenoids changes significantly upon exposure of *Gloeocapsa* sp. to UV-A radiation. The differential expression of these carotenoids by UV radiation leads to a question of the regulatory mechanisms involved in their biosynthesis. One of the peaks identified from the Gloeo-0311-UV extract was also found to have the characteristic UV profile of a carotenoid (**Figure A.2 D**). This peak had three λ_{max} = 448, 473, and 504 with a mass of *m/z* 567 as is typical for many carotenoids. This finding indicates that *Gloeocapsa* sp. is responding to UV-A radiation through the production of different carotenoids than are present under non-UV radiation conditions. The production of these carotenoids suggests that the biosynthetic genes responsible for their biosynthesis are also being upregulated.

Regulation of carotenoid biosynthesis has been shown in *Chlamydomonas reinhardtti*, a unicellular alga, to be controlled at a transcriptional level by light. In particular, a blue-light photoreceptor was shown to play a role in the response of two genes involved in the biosynthesis of carotenoids, phytoene synthase and phytoene desaturase (Bohne and Linden, 2002). The results obtained in this study suggest that cyanobacteria may also have a specific light induced regulatory mechanism involved in the biosynthesis of carotenoids. Further studies to understand the regulatory mechanism responsible for controlling the differential expression of carotenoids in cyanobacteria will provide valuable insight into the light regulated signaling mechanisms involved in this process.

Conclusions

Cyanobacteria are known to survive in environments of high light intensity including exposure to significant levels of UV radiation (Ehling-Shulz and Scherer, 1999). In this study, *Gloeocapsa* sp. is shown to produce three pigmented carotenoids in response to UV-A radiation. Carotenoids are important in nutrition and health as a source of vitamin A and for protection against disease due to their antioxidant activities (Britton *et al*., 1995). The upregulation of three carotenoids upon exposure to UV radiation suggests that *Gloeocapsa* sp. has a specific light regulatory mechanism signaling the biosynthesis of carotenoids in response to UV-A radiation. A better understanding of this mechanism may lead to the discovery of a UV inducible

signaling mechanism that could play a role in the commercial development of carotenoids for human health and nutrition.

Acknowledgements

I would like to acknowledge the late Dr. Ralph Lewin for providing the *Gloeocapsa* sp. cyanobacterial cells and for his insights into the pigmentation changes in this organism. Dr. Lewin was responsible for many advancements in the understanding of cyanobacteria throughout his career and will be remembered for his contributions to this field. I would also like to acknowledge R. Cameron Coates for the initial culturing of this organism.

References

Benton, M.J. (2009) The red queen and the court jester: species diversity and the role of biotic and abiotic factors through time. *Science*. 323: 728-732.

Björn, L.O. (2007) Stratospheric ozone, ultraviolet radiation, and cryptogams. *Biological Conservation*. 135: 326-333.

Bohne, F. and Linden, H. (2002) Regulation of carotenoid biosynthesis genes in response to light in *Chlamydomonas reinhard*tii. *Biochemica et Biophysica Acta*. 1579: 26-34.

Britton, G. (1995) UV/Visible Spectroscopy. In *Carotenoids*. Britton, G., Liaaen-Jensen, S., and Pfander H. eds. Birkhäuser Verlag. Switzerland. 2A: 13-62.

Britton, G., Liaaen-Jensen, S., and Pfander, H. (1995) Carotenoids today and challenges for the future. In *Carotenoids*. Britton, G., Liaaen-Jensen, S., and Pfander H. eds. Birkhäuser Verlag. Switzerland. 1A: 13-20.

Buckley, C.E. and Houghton, J.A. (1976) A study of the effects of near UV radiation on the pigmentation of the blue-green alga *Gloeocapsa alpicola*. *Arch. Microbiol*. 107: 93-97.

Castenolz, R.W. and Garcia-Pichel, F. (2000) Cyanobacterial responses to UV radiation. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M. Kluwer Academic Publishers: Netherlands, 591- 611.

Dismukes, G.C., Klimov, V.V., Baranov, S.V., Kozlov, Y.N., DasGupta, J., and Tyryshkin, A. (2001) The origin of atmospheric oxygen on Earth: The innovation of oxygenic photosynthesis. *Proc. Nat. Acad. Sci*. 98: 2170-2175.

Ehling-Schulz, M. and Scherer, S. (1999) UV protection in cyanobacteria. *European Journal of Phycology*. 34, 329-338.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 281: 237-240.

Glazer, A.N. (1977) Structure and molecular organization of the photosynthetic accessory pigments of cyanobacteria and red algae. *Mol. Cell. Biochem*. 18: 126-140.

Hirschberg, J. and Chamovitz, D. (1994) Carotenoids in cyanobacteria. In *The Molecular Biology of Cyanobacteria*. Bryant, D.A. Kluwer Academic Publishers, The Netherlands. 559-579.

He, Y.Y. and Häder, D.P. (2002) UV-B-induced formation of reactive oxygen species and oxidative damage of the cyanobacterium *Anabaena* sp.: protective effects of ascorbic acid and N-acetyl-L-cysteine. *J. Photochem. Photobiol. B*. 66: 115-124.

Lakatos, M., Bilger, W., and Büdel, B. (2001) Carotenoid content of terrestrial cyanobacteria: response to natural light conditions in open rock habitats in Venezuela. *Eur. J. Phycol*. 36: 367-375.

Lewin, R.A. (2006) Black algae. *J. Appl. Phycol*. 18: 699-702.

Mathews, C.K., van Holde, K.E., and Ahern, K.G. (2000) Photosynthesis. *Biochemistry* 3rd edition. Addison Wesley Longman, Inc: USA., 601.