THE PHAEOCHROMYCINS FROM *STREPTOMYCES* STRAIN LL-P018: FROM TAXONOMY TO NOVELTIES OF BIOSYNTHESIS

by

FRANK V. RITACCO

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Microbiology and Molecular Genetics

written under the direction of

Dr. Douglas E. Eveleigh

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2007

ABSTRACT OF THE DISSERTATION

The phaeochromycins from *Streptomyces* strain LL-P018: from taxonomy to novelties of biosynthesis

By FRANK V. RITACCO

Dissertation Director:
Dr. Douglas E. Eveleigh

The phaeochromycins are a newly discovered family of aromatic polyketides produced by *Streptomyces* strain LL-P018. The novel phaeochromycins A and C are of medical interest as they have inhibitory activity against MK-2, a kinase involved in the inflammatory response. The phaeochromycins are structurally similar to actinomycete secondary shunt metabolites derived from the actinorhodin and enterocin biosynthetic pathways. Because these phaeochromycins show promise as potential anti-inflammatory therapeutic agents, further knowledge of their biosynthesis and the taxonomy of the producing organism was sought. Strain LL-P018 was identified taxonomically using a combination of traditional and molecular techniques. Analyses of morphology, physiology, 16S ribosomal RNA (16SrRNA) sequence, and ribosomal polymerase β-subunit (*rpoB*) gene sequence were used to identify LL-P018 as a strain of *Streptomyces phaeochromogenes*. New methods for strain comparison by combining genetic fingerprints and metabolite profiles in a single comparison were developed to assess strain variation between strain LL-P018 and closely related organisms. This approach also clarified relationships between multiple "types" and other published strains of *S.*

phaeochromogenes and *Streptomyces ederensis*. Alnumycin, an additional aromatic polyketide which has structural similarities to the phaeochromycins, was produced by strain LL-P018 in liquid culture media. Genomic analysis of strain LL-P018 revealed the presence of a type II polyketide synthase gene cluster. Gene disruption experiments determined this pathway to be responsible for both phaeochromycin and alnumycin biosynthesis, suggesting that the phaeochromycins may be intermediates or shunt products in alnumycin biosynthesis.

This thesis describes a novel taxonomic approach to classification which integrates data from prior genetic and metabolic assessments into a single combined comparison. The *S. phaeochromogenes* type strain is defined, and taxonomic status of the species *S. ederensis* is clarified. New insights of phaeochromycin biosynthesis are revealed by cultural and genetic studies of *Streptomyces* strain LL-P018.

Preface

This thesis describes the result of a unique and fortuitous collaboration between my employer, Wyeth Research, and Rutgers University, the institution where I received both my Bachelor's and Master's degrees. Without the support and encouragement of my supervisors and advisors at both of these institutions, this project would not have been possible, and the opportunity for me to pursue a Ph.D. at this stage of my life could not have been realized.

This encouragement came first from Dr. Douglas Eveleigh, who was my undergraduate microbiology professor as well as a mentor and friend during my graduate studies. Doug convinced me not only of my scientific ability to achieve the degree, but also of his willingness to shepherd me through the process. As I had already completed several years of the true Rutgers experience, Doug found a manageable way for me to continue my studies within the framework of the graduate program at Rutgers, while still working full-time at Wyeth Research. Without his experience and guidance this would not have been possible.

Further support came from my former supervisor (now happily retired) at Wyeth Research, Dr. Michael Greenstein. Michael has had a tremendous influence on my scientific career, and advised me that completion of my doctorate would be rewarding enough that several tough years of juggling work, school, and family life would prove

worthwhile. Although he is no longer at Wyeth, we have kept in touch and I still appreciate his continuing advice and encouragement.

Countless other individuals have helped me tremendously in my studies. My supervisors at Wyeth, Dr. Valerie Bernan, Dr. Leonard McDonald, Dr. Frank Koehn and Dr. Guy Carter have allowed me the flexibility to mix school and work. They also provided financial support from Wyeth, as well as technical and scientific advice. Many of my colleagues at Wyeth, including Dr. Edmund Graziani, Brad Haltli, Dr. Kerry Kulowski, and Laurie Barbieri made important contributions to this project and to my education by teaching me the methods and science that fueled this research. Dr. Stanley Katz and Dr. Jerry Kukor, my thesis committee members at Rutgers, have listened patiently to my research presentations and have advised me not only scientifically, but also practically, regarding the important matter of actually completing a thesis within a reasonable timeframe. Dr. David Labeda, curator of the Actinobacterial Collection of the USDA Agricultural Research Service Culture Collection, and a Chair on the International Committee on Systematics of Prokaryotes (Chair of the Subcommittee on the Taxonomy of the *Streptomycetaceae*) has also been of great assistance by providing strains for evaluation as well as advice and protocols for taxonomic studies.

Lastly, and most importantly, my family has made tremendous sacrifices so that I could pursue this opportunity. They have been infinitely patient, supportive and motivating. They are the reason that I returned to school, and they are the reason that I am nearing completion. I am forever grateful.

Dedication

This thesis is dedicated to my wife, Gretchen, and to my children, Cecilia and Matthew. Thank you for supporting me, encouraging me, keeping me sane, and most of all, for putting up with me.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Illustrations

List of Appendices

Chapter 1

Introduction

Actinomycetes have long been studied as potential sources of novel natural products with useful pharmacological activities (Baltz, 1997, 2006; Challis & Hopwood, 2003; Chater & Bibb, 1997; Hansen, 1997; Jack *et al.*, 1997; Mankelow & Neilan, 2000; Nicas & Cooper, 1997; Piepersberg & Distler, 1997; Schatz *et al.*, 1944; Sehgal *et al.*, 1975; Strohl, 1997; Waksman & Woodruff, 1940, 1942; Walsh, 2003, 2004; Wink, 2002). Species of the genus *Streptomyces*, which are abundant in soil, are particularly bounteous producers of bioactive secondary metabolites. As such, streptomycetes have been a major component in pharmaceutical screening programs, in which isolates are cultured and then analyzed for the production of metabolites with distinctive biological activity, pigmentation, or chemical spectra. Numerous *Streptomyces* metabolites with desirable pharmacological properties have been discovered in this manner, resulting in the development of breakthrough antibacterial, antifungal, anticancer, antiparasitic, and immunosuppressant therapeutic agents for the treatment of human disease (see Chapter 2, Table 2.1, page 37).

The discovery of a promising new microbial metabolite is typically followed by a thorough study of the taxonomy and metabolism of the producing organism. Its taxonomy reveals evolutionary perspective and clarifies relationships to closely related taxa (Anderson & Wellington, 2001; Goodfellow, 1989). Consideration of the metabolite

and analogous compounds provides further evolutionary and metabolic insight. Culture conditions for production of the desired metabolite are optimized, allowing both improved yields and understanding of the regulation of biosynthesis (Demain & Vaishnav, 2004; Sanchez & Demain, 2002). Analysis of the genes encoding the biosynthetic enzymes takes the investigation further, providing a clearer view of the nature of the biosynthetic mechanism, and revealing possible approaches to increasing product yield or generating novel analogs through genetic engineering (August *et al.*, 1999; Walsh, 2003).

Central to this thesis, *Streptomyces* strain LL-P018 was originally isolated from the bank of the Aa river in Westerenger, Germany (Gerhard Schlingmann, Wyeth Research, personal communication), and became part of the culture collection at Wyeth Research in Pearl River, NY (Graziani *et al.*, 2005; Ritacco & Eveleigh, 2007). Wyeth's natural products research program generates a library of microbial fermentation extracts, which are screened for activity against a range of therapeutic targets *via* a high throughput-screening (HTS) program. Strain LL-P018 was cultured in a variety of liquid media, and yielded culture extracts which were of interest for their inhibition of MAP kinase-activated protein kinase 2 (MAPKAP-2, MK-2), an enzyme involved in inflammation. MK-2 activity is required for the production of tumor necrosis factor (TNF-α) and other inflammatory cytokines during the inflammatory response (Kotlyarov *et al.*, 1999). Because elimination of MK-2 activity has been shown to dramatically reduce TNF-α production, small molecule inhibitors of MK-2 are of interest as potential anti-inflammatory drug candidates (Gaestel, 2006).

Culture extracts from strain LL-P018 were analyzed by high performance liquid chromatography coupled with mass spectrometry (HPLC-MS). Active metabolites were purified from large-scale fermentations *via* bioassay-guided fractionation, and their structures were determined by nuclear magnetic resonance (NMR) spectroscopy. The active compounds were identified as a new family of aromatic polyketides, and were named the phaeochromycins (Graziani *et al.*, 2005). The phaeochromycins (Figure 1.1) are analogous to partially cyclized shunt products of the actinorhodin and enterocin biosynthetic pathways. This similarity suggested that the phaeochromycins themselves may be partially cyclized intermediates or shunt products derived from a biosynthetic pathway for which the fully cyclized end product had not yet been identified (Graziani *et al.*, 2005; Kalaitzis & Moore, 2004; McDaniel *et al.*, 1994a, b; Xiang *et al.*, 2002). Because these compounds have importance as potential anti-inflammatory therapeutic agents, further understanding of their biosynthesis and the taxonomy of strain LL-P018 was sought.

Strain LL-P018 was taxonomically characterized using a variety of morphological, physiological, and molecular biological analyses. 16S ribosomal RNA (16SrRNA) and ribosomal polymerase β-subunit (*rpoB*) gene sequencing were used along with classical taxonomic methods (such as morphological observation, carbohydrate utilization, sporulation and pigment production) to identify the genus and species of this organism. However, more sensitive methods were required in order to determine the uniqueness of

strain LL-P018 with regard to its production of novel secondary metabolites (Anderson & Wellington, 2001; Anzai *et al.*, 1997; Ritacco *et al.*, 2003).

Strain LL-P018 was compared to fifteen closely related strains from major culture collections by genetic and chemical fingerprinting methods. Genetic fingerprinting was performed using the RiboPrinter®, an automated ribotyping instrument capable of discriminating between strains of *Streptomyces* (Pfaller *et al.*, 1996; Ritacco *et al.*, 2003). For comparison of chemical fingerprints, all strains were grown in multiple liquid media and analyzed by HPLC-MS. BioNumerics® software (Applied Maths, Austin, TX) was used to integrate RiboPrint® patterns and multiple HPLC chromatograms in a single cluster analysis, allowing simultaneous comparison of molecular typing data and metabolite profiles.

With the objectives of identifying new phaeochromycins, and possibly the fully elaborated end product of phaeochromycin biosynthesis, culture extracts from strain LL-P018 were analyzed for phaeochromycins and related metabolites. Compounds of interest were purified from one liter cultures by reverse-phase preparative HPLC, and structures were identified using NMR spectroscopy.

In order to identify the presence of genes encoding phaeochromycin biosynthesis, the genome of strain LL-P018 was examined for type II polyketide synthase (PKS) genes through use of degenerate primers combined with the use of the polymerase chain reaction (PCR). A putative ketosynthase gene was identified, and then used to probe a

cosmid library. By this method, a portion of a type II PKS pathway from strain LL-P018 was identified. Gene functions were proposed based on BLAST analysis (Altschul *et al.*, 1997), and a knockout construct was developed and used to confirm the role of these genes in the biosynthesis of the phaeochromycins and other metabolites.

It was hypothesized that as the phaeochromycins are a new family of bioactive polyketides, which represent potential novelties of taxonomy and biosynthesis, that further analysis of strain LL-P018 would gradually unravel phaeochromycin biosynthesis while leading to the discovery of new phaeochromycin analogs or related compounds. This hypothesis could be supported through analysis of phaeochromycin biosynthetic genes, potentially leading to the identification of a novel biosynthetic pathway. Furthermore, taxonomic identification of strain LL-P018 and comparative evaluation of closely related strains will provide useful new evolutionary perspective on the production of these and other biologically active polyketides from the streptomycetes.

The aims of this thesis are:

1. Taxonomic identification of *Streptomyces* strain LL-P018, and comparison to closely related strains.
2. Screening for new phaeochromycin analogs or related compounds, to enhance our understanding of phaeochromycin biosynthesis.

3. Identification of genes involved in phaeochromycin biosynthesis, which will ultimately allow analysis of the biosynthetic gene cluster and enable the production of new molecules *via* genetic engineering.
4. Illumination of taxonomic and biosynthetic novelties, which will broaden and strengthen our understanding of *Streptomyces*.

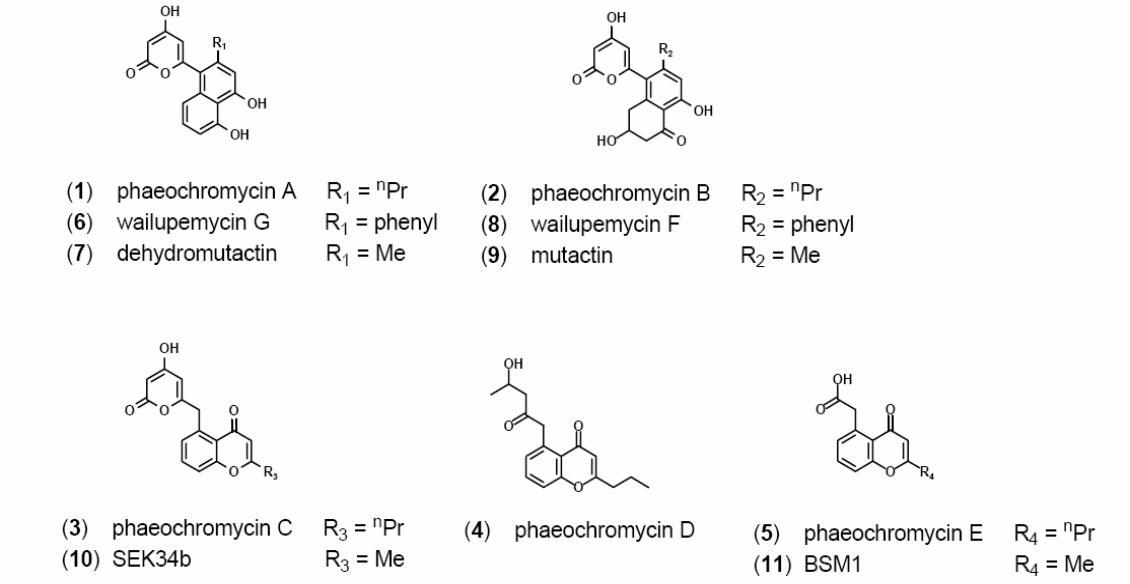


Figure 1.1. Structures of phaeochromycins A-E (1 to 5) and related metabolites. The phaeochromycins are analogous to shunt metabolites from the actinorhodin and enterocin biosynthetic pathways, namely wailupemycin G (6), dehydromutactin (7), wailupemycin F (8), mutactin (9), SEK34b (10), and BSM1 (11). The phaeochromycins differ from these analogous metabolites by the presence of a C4 starter unit, presumably derived from butyrate. Me = methyl, ^{n}Pr = n-propyl (Graziani *et al.*, 2005; Hopwood, 1997; Kalaitzis & Moore, 2004; McDaniel *et al.*, 1994a, b; Xiang *et al.*, 2002).

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Anderson, A. S. & Wellington, E. M. H. (2001). The taxonomy of *Streptomyces* and related genera. *Int J Syst Evol Microbiol* **51**, 797-814.

Anzai, Y., Okuda, T. & Watanabe, J. (1997). Application of the random amplified polymorphic DNA using the polymerase chain reaction for efficient elimination of duplicate strains in microbial screening. II. Actinomycetes. *J Antibiot* **50**, C-1.

August, P. R., Yu, T.-W. & Floss, H. G. (1999). Molecular biological aspects of antibiotic biosynthesis. In *Drug Discovery from Nature*, pp. 215-232. Ed. S. Grabley & R. Thiericke. Springer-Verlag, Berlin, Germany.

Baltz, R. H. (1997). Lipopeptide antibiotics produced by *Streptomyces roseosporus* and *Streptomyces fradiae*. In *Drugs and the Pharmaceutical Sciences*, pp. 415-435. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Baltz, R. H. (2006). Molecular engineering approaches to peptide, polyketide and other antibiotics. *Nat Biotechnol* **24**, 1533-1540.

Challis, G. L. & Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci USA* **100 (Suppl 2)**, 14555-14561.

Chater, K. F. & Bibb, M. J. (1997). Regulation of bacterial antibiotic production. In *Biotechnology (2nd Edition)*, pp. 57-105. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Demain, A. L. & Vaishnav, P. (2004). Secondary metabolism in microbes and its control by phosphate and metals. *SIM News* **54**, 104-113.

Gaestel, M. (2006). MAPKAP kinases - MKs - two's company, three's a crowd. *Nat Rev Mol Cell Biol* **7**, 120-130.

Goodfellow, M. (1989). The Actinomycetes I. Suprageneric classification of actinomycetes. In *Bergey's Manual of Systematic Bacteriology*, pp. 2333-2339. Ed. S. T. Williams, M. E. Sharpe & J. G. Holt. Williams & Wilkins, Baltimore, MD.

Graziani, E. I., Ritacco, F. V., Bernan, V. S. & Telliez, J. B. (2005). Phaeochromycins A-E, anti-inflammatory polyketides isolated from the soil actinomycete *Streptomyces phaeochromogenes* LL-P018. *J Nat Prod* **68**, 1262-1265.

Hansen, J. N. (1997). Nisin and related antimicrobial peptides. In *Drugs and the Pharmaceutical Sciences*, pp. 437-470. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Hopwood, D. A. (1997). Genetic contributions to understanding polyketide synthases. *Chem Rev* **97**, 2465-2497.

Jack, R., Gotz, F. & Jung, G. (1997). Lantibiotics. In *Biotechnology (2nd Edition)*, pp. 323-368. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Kalaitzis, J. A. & Moore, B. S. (2004). Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* **67**, 1419-1422.

Kotlyarov, A., Neininger, A., Schubert, C., Eckert, R., Birchmeier, C., Volk, H. D. & Gaestel, M. (1999). MAPKAP kinase 2 is essential for LPS-induced TNF-alpha biosynthesis. *Nat Cell Biol* **1**, 94-97.

Mankelow, D. P. & Neilan, B. A. (2000). Non-ribosomal peptide antibiotics. *Exp Op Therap Pat* **10**, 1583-1591.

McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A. & Khosla, C. (1994a). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci USA* **91**, 11542-11546.

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1994b). Engineered biosynthesis of novel polyketides: actVII and actIV genes encode aromatase and cyclase enzymes, respectively. *J Am Chem Soc* **116**, 10855-10859.

Nicas, T. I. & Cooper, R. D. G. (1997). Vancomycin and other glycopeptides. In *Drugs and the Pharmaceutical Sciences*, pp. 363-392. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Pfaller, M. A., Wendt, C., Hollis, R. J., Wenzel, R. P., Fritschel, S. J., Neubauer, J. J. & Herwaldt, L. A. (1996). Comparative evaluation of an automated ribotyping system versus pulsed-field gel electrophoresis for epidemiological typing of clinical isolates of *Escherichia coli* and *Pseudomonas aeruginosa* from patients with recurrent gram-negative bacteremia. *Diagn Microbiol Infect Dis* **25**, 1-8.

Piepersberg, W. & Distler, J. (1997). Aminoglycosides and sugar components in other secondary metabolites. In *Biotechnology (2nd Edition)*, pp. 397-488. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Ritacco, F. V., Haltli, B., Janso, J. E., Greenstein, M. & Bernan, V. S. (2003). Dereplication of *Streptomyces* soil isolates and detection of specific biosynthetic genes using an automated ribotyping instrument. *J Ind Microbiol Biotechnol* **30**, 472-479.

Ritacco, F. V. & Eveleigh, D. E. (2007). Molecular and phenotypic comparison of phaeochromycin-producing strains of *Streptomyces phaeochromogenes* and *Streptomyces ederensis*. *Manuscript in preparation*.

Sanchez, S. & Demain, A. L. (2002). Metabolic regulation of fermentation processes. *Enzyme Microb Technol* **31**, 895-906.

Schatz, A., Bugie, E. & Waksman, S. A. (1944). Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc Soc Exp Biol Med* **55**, 66-69.

Sehgal, S. N., Baker, H. & Vezina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. *J Antibiot (Tokyo)* **28**, 727-732.

Strohl, W. R. (1997). Industrial antibiotics: today and the future. In *Drugs and the Pharmaceutical Sciences*, pp. 1-47. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Waksman, S. A. & Woodruff, H. B. (1940). Bacteriostatic and bactericidal substances produced by a soil *actinomyces*. *Proc Soc Exp Biol Med* **45**, 609-614.

Waksman, S. A. & Woodruff, H. B. (1942). Streptothricin, a new selective bacteriostatic and bactericidal agent, particularly active against gram-negative bacteria. *Proc Soc Exp Biol Med* **49**, 207-210.

Walsh, C. T. (2003). *Antibiotics: actions, origins, resistance*.ASM Press, Washington, D.C.

Walsh, C. T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805-1810.

Wink, J. (2002). *The Actinomycetales: an order in the class of Actinobacteria important to the pharmaceutical industry- electronic manual*.CD-Rom Provided by Dr. Joachim Wink, Aventis Pharma Deutschland GmbH.

Xiang, L., Kalaitzis, J. A., Nilsen, G., Chen, L. & Moore, B. S. (2002). Mutational analysis of the enterocin favorskii biosynthetic rearrangement. *Org Lett* **4**, 957-960.

Chapter 2

Review of the Literature

The genus *Streptomyces*, first described in 1943 by Waksman and Henrici, consists of aerobic, gram-positive bacteria, which exhibit mycelial growth. They are extremely abundant in soil. This genus belongs to the order *Actinomycetales* within the class *Actinobacteria* (Anderson & Wellington, 2001; Stackebrandt et al., 1997). Because streptomycetes and actinomycetes in general are prolific producers of diverse antibiotics and other biologically active metabolites, they have been studied extensively over the past 50 years. These microorganisms have been the source of a tremendous number of medically important and highly valuable pharmaceutical products including antibacterial, antifungal, antiparasitic, anticancer and immunosuppressant drugs (Table 2.1) (Baltz, 1997, 2006; Challis & Hopwood, 2003; Chater & Bibb, 1997; Hansen, 1997; Jack *et al.*, 1997; Mankelow & Neilan, 2000; Nicas & Cooper, 1997; Piepersberg & Distler, 1997; Schatz *et al.*, 1944; Sehgal *et al.*, 1975; Strohl, 1997; Waksman & Woodruff, 1940, 1942; Walsh, 2003, 2004; Wink, 2002).

Since the discoveries of actinomycin (Waksman & Woodruff, 1940), streptothricin (Waksman & Woodruff, 1942) and streptomycin (Schatz *et al.*, 1944) at Rutgers University demonstrated the potential for soil bacteriology to yield such valuable pharmaceuticals, academic and industrial institutions have embraced the science of *Streptomyces* biology by engaging in the field of natural products research (Omura, 1992; Strohl, 1997; Walsh, 2003). In the classical screening approach, microorganisms

(primarily actinomycetes and filamentous fungi) are isolated from soil, plants, marine sediments, or other environmental samples (Hunter-Cevera & Belt, 1999). These cultures are compared morphologically and taxonomically, catalogued, and a diverse collection of industrially significant microorganisms is built. The collected microbes are each grown under various nutritional and environmental conditions known to promote the production of secondary metabolites. Culture broths are analyzed chemically and biologically for the presence of bioactive metabolites. This avenue of research has yielded many of the valuable, *Streptomyces*-derived pharmaceutical products mentioned previously (Table 2.1).

In the pharmaceutical industry, natural products research groups typically define a "lead" compound as one with desirable chemical novelty combined with biological activity as identified by a particular screen or assay. Lead compounds are evaluated in great detail for their pharmacological properties as well as feasibility of large-scale manufacturing. For a target compound that is produced by microbial fermentation, the taxonomy and metabolism of the producing organism are studied in great detail, in order to elucidate the biosynthesis of the target compound. Such analyses yield greater understanding and can therefore lead to more efficient synthesis.

Taxonomy

Taxonomic evaluation of streptomycetes requires a polyphasic approach incorporating traditional morphological and phenotypic methods along with modern molecular

biological techniques (Anderson & Wellington, 2001). In the 1960s, a set of standardized methods for the description of *Streptomyces* cultures for taxonomic purposes was established as part of the International *Streptomyces* Project (Shirling & Gottlieb, 1966). Using these methods, a *Streptomyces* species can be described on the basis of spore color, spore chain and spore surface morphologies, coloration of substrate and aerial mycelia, production of soluble pigment, and carbohydrate utilization profile. The authors defined standardized media and methods in order to ensure descriptive consistency. Keys were developed that allow rapid comparison of an unknown *Streptomyces* isolate to previously described species using these established criteria (Nonomura, 1974).

Chemotaxonomic methods are those that determine relatedness or taxonomic grouping based on cellular phenotype. Typically, these methods are based on analyses of the chemical makeup of the bacterial cell. These chemotaxonomic techniques are important for distinguishing the genus *Streptomyces* from other genera, but are not very useful for identification of species within the genus (Anderson & Wellington, 2001).

Determination of the diaminopimelic acid isomer in the cell wall can be helpful in identifying the genus of an unknown actinomycete (Bernan *et al.*, 1994; Goodfellow, 1989; Nagasawa *et al.*, 1993). Members of the genus *Streptomyces* possess a type I cell wall (Lechevalier & Lechevalier, 1970), which by definition contains glycine and the LL-isomer of diaminopimelic acid (LL-DAP) within the peptidoglycan. Many other genera of actinomycetes (*Micromonospora, Nocardia*) tend to contain the *meso*-DAP isomer (Goodfellow, 1989).

Analysis of whole cell fatty acid composition has been used to identify members of the genus *Streptomyces,* which typically possess a mixture of saturated straight-chain, iso-branched and anteiso-branched chain fatty acids, ranging in carbon-chain length from about 14 to 18 atoms (Anderson & Wellington, 2001; Goodfellow *et al.*, 1992; Lechevalier, 1977). Fatty acid methyl ester (FAME) analysis has been used to identify and compare streptomycetes below the genus level (Saddler *et al.*, 1987). Variations in measured fatty acid content resulting from growth conditions and analytical methods make this technique insufficient for accurately determining species identification (Anderson & Wellington, 2001; Saddler *et al.*, 1986). However, this method can be effective in distinguishing between unknown environmental isolates (Ritacco *et al.*, 2003; Saddler *et al.*, 1987).

Other chemotaxonomic methods include analysis of cellular menaquinones, which are involved in electron transport (Goodfellow, 1989; Moss & Guerrant, 1983), and the analysis of whole cell sugars and phospholipids (Minnikin *et al.*, 1984; Lechevalier, 1977; Lechevalier and Lechevalier, 1970; Saddler *et al.*, 1991). The *Streptomyces* cell typically contains MK-9 menaquinones (with an isoprenoid chain nine carbons in length), and lacks a diagnostic pattern of sugars, while the predominant phospholipid is usually phosphatidylethanolamine (Goodfellow, 1989). Thus these methods are often used for genus identification or discrimination between genera, but are not species-specific.

Advances in molecular biology have revolutionized the science of microbial taxonomy. Methods involving nucleic acid fingerprinting and gene sequence comparisons have become routine practice in species identification. DNA-DNA hybridization has long been used to determine genetic relatedness between *Streptomyces* species, by measuring the percent homology of total chromosomal DNA between strains (Labeda, 1992). However, this method is laborious and time-consuming, and as such, newer approaches such as gene sequencing and genetic fingerprinting are often used to complement this approach (Hatano, *et al.*, 2003; Lanoot, *et al.*, 2005).

Analysis of the 16S ribosomal RNA gene sequence (16S rRNA), originally championed by Carl Woese as a means for understanding microbial evolution and relationships (Woese, 1987), has become a universal tool in streptomycete taxonomy. Because this gene is highly conserved among bacteria, yet with distinctive variations, 16S rRNA sequence comparisons are widely used in genus, species, and strain identification (Anderson & Wellington, 2001; Stackebrandt *et al.*, 1991, 1992). The complete 16S gene sequence can be easily amplified using the polymerase chain reaction (PCR) and primers derived from highly conserved regions at both ends of the gene (Reysenbach *et al.*, 1994; Stackebrandt *et al.*, 1991), while automated sequencing equipment can efficiently sequence the PCR product.

Three variable regions, labeled alpha, beta, and gamma, have been identified within the 16S rRNA sequence of *Streptomyces* (Stackebrandt *et al.*, 1991). The gamma region (located around nucleotide position 190) is the most variable of the three, and analysis of

this region can yield species-specific signatures useful for taxonomic identification and comparison (Stackebrandt *et al.*, 1992). Sequencing of the entire 16S rRNA gene (approximately 1500 base pairs) and comparison with the GenBank database is now fundamental to the investigation of the taxonomy of a streptomycete. However, this database is still incomplete, and for many unknown environmental isolates a close match may not be available (Anderson & Wellington, 2001).

Overall, while 16S rRNA gene sequence comparison may be the most powerful tool for phylogenetic analysis of streptomycetes, a polyphasic approach combining traditional methods is still necessary for species identification.

In addition to 16S rRNA, other gene sequences have been used for taxonomic identification of *Streptomyces*. RNA polymerase β-subunit (*rpoB*) partial gene sequencing has been shown to be an effective tool for *Streptomyces* taxonomy, with certain advantages over 16S rRNA gene sequencing (Kim, *et al.*, 2004). The *gyrB* gene, which encodes the B subunit of DNA gyrase, has also been successfully used for species level comparison of *Streptomyces* species (Hatano, *et al.*, 2003). These methods complement other genetic comparisons such as 16S rRNA gene sequence and DNA-DNA hybridization, and can be very useful in clarifying relationships between *Streptomyces* species. As the number of *Streptomyces rpoB* and *gyrB* sequences in public databases increases, these sequences will become even more valuable in the identification of unknown isolates and comparison to validly described species.

Genetic fingerprinting methods, such as restriction fragment length polymorphisms (RFLP), randomly amplified polymorphic DNA (RAPD), and ribotyping, have been widely utilized in order to discriminate between closely related *Streptomyces* strains, or to identify strain diversity in a group of environmental isolates (Anderson & Wellington, 2001; Anzai *et al.*, 1997; Lanoot, *et al.*, 2005; Ritacco *et al.*, 2003). These molecular typing methods provide a greater level of sensitivity than 16S rRNA gene sequence analysis, and are suitable for discrimination of subspecies or strain variation below the species level. RFLP involves digestion of chromosomal DNA with a restriction endonuclease, followed by analysis of fragment sizes by pulsed-field gel electrophoresis (Anderson & Wellington, 2001). The pattern of DNA fragments visualized as bands following electrophoresis serves as a fingerprint, which can is used to compare strains. Similar band pattern "fingerprints" are generated by RAPD as well as ribotyping, but by different means. RAPD uses PCR with random, short primers to generate a pattern of bands amplified from the target organism's genomic DNA (Anzai *et al.*, 1997), while ribotyping basically utilizes Southern blotting to identify fragments of digested chromosomal DNA that hybridize with a ribosomal probe (Ritacco *et al.*, 2003). Each of these methods is effective for detecting strain variation without the need for sequencing, as long as consistency is maintained in the amplification, hybridization and electrophoresis conditions used.

Production of actinomycete secondary metabolites

In order to achieve effective production of desired metabolites by actinomycetes, culture

media are designed which provide the microorganisms with substrates for generation of metabolic energy, as well as materials to be used in the synthesis of cellular material and metabolic products (Dahod, 1999; Miller & Churchill, 1986; Zabriskie *et al.*, 1980). Many industrially significant microbially produced compounds are classified as secondary metabolites. Secondary metabolites are those products of microbial metabolism that are not essential for the growth of the producing organism, but probably play some role in the survival of the organism in its natural habitat, possibly by providing competitive advantage in this environment (Demain & Vaishnav, 2004; Strohl, 1997). Natural product antibiotics and many other pharmaceutical products derived from streptomycetes are secondary metabolites (Strohl, 1997; Walsh, 2003).

Secondary metabolites fall into distinct chemical groups. In bacteria, the predominant secondary metabolites are polyketides, oligoglycosidic (sugar-derived) compounds, and peptide (amino acid-derived) compounds. Polyketide structures are generated by sequential reactions linking small acyl-CoA precursors *via* decarboxylating condensation, in a manner which resembles fatty acid synthesis in bacteria and mammals (Pfeifer & Khosla, 2001; von Dohren & Grafe, 1997). Polyketide structures are highly diverse (Table 2.2) and include several subgroups, including the macrolides (such as erythromycin), polycyclic and aromatic compounds (actinorhodin, tetracycline, doxorubicin), polyenes (nystatin) and polyethers (monensin) (Strohl, 1997). Oligoglycosides are derived from sugars activated into nucleoside sugars, which are subsequently linked to other such activated sugars (von Dohren & Grafe, 1997). The aminoglycosides, such as streptomycin (Table 2.1), are among the most studied and

commercially successful members of this group. These compounds contain amino nitrogen present either in saccharides, or in cyclitol derivatives (such as amino-inositol in streptomycin) or both (Piepersberg & Distler, 1997). Peptide antibiotics and other amino acid-derived secondary metabolites are diverse in terms of structure and biosynthetic origins (Table 2.3). In addition to basic linear and cyclic peptides, there exist various modified and substituted varieties: the β-lactams (penicillin, cephalosporins) in which the tripeptide δ-(L-α-aminoadipyl-L-cysteinyl-D-valine) is cyclized to form the β-lactam ring that is critical for antimicrobial activity (Paradkar *et al.*, 1997); the lipopeptides (such as daptomycin) in which the peptide is substituted with fatty acids (Baltz, 1997; von Dohren & Grafe, 1997); the glycopeptides (such as vancomycin and teicoplanin) all of which contain a similar heptapeptide backbone structure, and some of which are substituted with sugar moieties (Hubbard & Walsh, 2003; Nicas & Cooper, 1997); and the ribosomally-synthesized lantibiotics (such as mersacidin) in which the original peptide produced by the ribosome undergoes post-translational modification such that the final product contains the unusual non-protein amino acids lanthionine and 3-methyl lanthionine (Hansen, 1997; Jack *et al.*, 1997). Many mixed-type compounds also exist, whose structures are derived from hybrid biosynthetic pathways containing aspects of several different biosynthetic mechanisms. For example, rapamycin is produced by a mixed polyketide synthase/ non-ribosomal peptide synthetase gene cluster. As a result, the structure of rapamycin contains a polyketide backbone that includes an incorporated amino acid residue (Aparicio *et al.*, 1996; Molnar *et al.*, 1996; Schwecke *et al.*, 1995).

Terpenoids, another type of secondary metabolite, are very common in plants, but are rarely produced by bacteria (von Dohren & Grafe, 1997). Biosynthesis of these compounds requires mevalonic acid derived from acetyl CoA as a precursor, extended by subsequent addition of isoprene units (Demain & Vaishnav, 2004).

Microbial culture media are designed to provide the microorganism with the nutritional elements required for optimal cellular growth and rapid synthesis of metabolic products. These media can be subsequently optimized for the production of specific secondary metabolites of interest. Actinomycete culture media typically contain carbohydrates as a source of carbon, as either sugars such as glucose and sucrose, or as polymers such as starch or dextrin. Carbohydrates may also be provided in the form of molasses (sucrose), which also acts as a good source of nitrogen, vitamins and inorganic components. Other carbon sources that can be used as medium components include oils, alcohols, and organic acids such as acetic acid (Dahod, 1999; Zabriskie *et al.*, 1980).

Nitrogen for the synthesis of cellular proteins, nucleic acids and nitrogen-containing metabolites is typically supplied in the form of crude protein hydrolysates, such as plant-derived peptones or yeast extract. For defined culture media, nitrogen is supplied in pure forms such as amino acids, urea, or inorganic ammonia as ammonium hydroxide (Dahod, 1999; Miller & Churchill, 1986). Media contain potassium phosphate (or sodium phosphate) and sulfates (such as magnesium sulfate) to promote cellular growth, while essential trace elements (mineral salts) are added as required for growth and metabolite production. Typical trace elements include calcium, cobalt, copper, iron, magnesium,

manganese, molybdenum, sodium, and zinc (Dahod, 1999; Miller & Churchill, 1986; Zabriskie *et al.*, 1980).

Frequently, secondary metabolites are produced at highest rates when the growth rate of the organism is relatively low, often after rapid growth has stopped (Chater & Bibb, 1997; von Dohren & Grafe, 1997). Production rates can often be improved by addressing deficiencies in nutritional requirements or by supplementing culture media with limiting precursors (Demain & Vaishnav, 2004). However, physio-chemical factors also play a most important role in the regulation of antibiotic biosynthesis.

Secondary metabolite biosynthesis has been proposed to have evolved as a means for microorganisms to adapt to changing conditions in their environment, and at times, to provide a competitive advantage (Chater & Bibb, 1997; von Dohren & Grafe, 1997). Soil bacteria are often subjected to great environmental stresses, and can survive and also develop at extremely slow growth rates under starvation conditions. In nature as well as under laboratory conditions, stress conditions such as nutrient deficiency induce complex differentiation in these organisms, including sporulation and secondary metabolism (Kieser *et al.*, 2000). One possible evolutionary reason for this is that during times of high nutrient availability and low environmental stress, microbial metabolism is geared towards rapid growth and biomass production, which would be hindered by the energy exhaustive and precursor-demanding biosynthesis of large, complex secondary metabolites (Chater & Bibb, 1997). Production of antibiotics and other bioactive secondary metabolites may therefore be reserved for more stressful circumstances when

growth rate is low and the competitive advantage provided by these compounds can aid in long-term survival. Thus it has been reasoned that many nutrients which are essential for rapid growth and biomass accumulation can have a negative effect on antibiotic production (Piepersberg & Distler, 1997; Sanchez & Demain, 2002).

Most nutrients can act as either positive or negative regulators of metabolite production. Carbon sources can often act as repressors of secondary metabolite production. In regulation by carbon catabolite repression, biosynthesis of secondary metabolites is repressed when the microorganism is utilizing a certain carbon source, usually the one that allows the most rapid growth (Kieser *et al.*, 2000; Sanchez & Demain, 2002). Thus in the presence of the optimal carbon source for growth, the utilization of other carbon sources is prevented. This is achieved either by feedback inhibition of the activity of enzymes involved in the assimilation of the secondary carbon source, or by feedback repression of transcription of the genes responsible for the synthesis of such enzymes (Chater & Bibb, 1997). Glucose, for example, following phosphorylation by glucose kinase, adversely affects the production of aminoglycoside antibiotics in *Streptomyces* by repressing the activity of biosynthetic enzymes (Piepersberg & Distler, 1997). Culture media often contain a mixture of readily utilized carbon sources, such as glucose, and slowly utilized carbon sources such as starches or oils. Cells grow rapidly on the optimal carbon source, accumulating biomass without producing secondary metabolites. Depletion of the "best" source of carbon triggers a metabolic shift wherein the organism begins utilizing the less favorable carbon source, and secondary metabolite production proceeds (Dahod, 1999; Sanchez & Demain, 2002).

In a similar fashion, nitrogen sources that favor rapid growth can repress bacterial secondary metabolite production. Analagously to carbon catabolite repression, readily utilized nitrogen sources that are optimal for growth, such as ammonium salts, can inhibit the utilization of alternate nitrogen sources in the medium and suppress antibiotic production. Culture media typically include more complex nitrogen sources, such as crude protein hydrolysates, based on effectiveness and cost (Dahod, 1999; Miller & Churchill, 1986; Sanchez & Demain, 2002).

Inorganic components, such as inorganic phosphate and metals, also can regulate secondary metabolite production by various mechanisms. Phosphate availability favors rapid growth, but can also delay secondary metabolite production, as seen in carbon and nitrogen repression. Additionally, biosynthetic enzymes that phosphorylate intermediates in a pathway, or which require phosphorylation or dephosphorylation themselves for activity, can be very sensitive to inorganic phosphate concentration (Piepersberg & Distler, 1997). Metals have regulatory effects on secondary metabolic pathways in many microorganisms. Copper, for example, is often helpful for stimulation of sporulation in *Streptomyces*, and improves antibiotic production in *S. griseus* and *S. coelicolor* (Demain & Vaishnav, 2004; Kieser & Hopwood, 1991; Ueda *et al.*, 1997). Magnesium enhances vancomycin production by *Amycolatopsis orientalis,* and iron deficiency has a positive effect on actinorhodin production by *S. coelicolor* (Coisne *et al.*, 1999).

An important factor that contributes to the regulation of secondary metabolism is population density. Many bacteria, including streptomycetes, produce small, diffusible autoregulators. In quorum sensing, these molecules accumulate in the environment to threshold levels that trigger the induction of genes involved in secondary metabolism and differentiation (Bassler & Losick, 2006; Kawabuchi *et al.*, 1997; Kieser *et al.*, 2000). streptomycetes produce the autoregulatory signaling γ-butyrolactones. This class of compounds share a common acylated lactone skeleton, but differ slightly in acyl chain length and minor substitutions to the acyl chain. These seemingly minor structural differences are responsible for the ligand specificity of these molecules, affecting their binding to various receptor proteins (Chater & Bibb, 1997). Well-studied examples include the A-factor from *S. griseus*, which controls the expression of all of the regulatory genes for streptomycin production and is required for morphological differentiation in this organism (Chater & Bibb, 1997; Distler *et al.*, 1992; Piepersberg & Distler, 1997), and the butanolides which induce virginiamycin production in *S. virginiae* (Kawabuchi *et al.*, 1997; Yamada *et al.*, 1997).

Molecular Biology of Secondary Metabolite Biosynthesis in *Streptomyces* and other Actinomycetes

Complementary to the traditional screening of actinomycetes such as streptomycetes for production of novel secondary metabolites, considerable research now focuses on the genes responsible for the biosynthesis of these important microbial products. Pioneering work over the past decade has led to the identification and sequencing of gene clusters

encoding the biosynthetic pathways for several different types of antibiotics, in particular the polyketides (Aparicio *et al.*, 1996; Cortes *et al.*, 1990; Donadio *et al.*, 1991; Fernandez-Moreno *et al.*, 1992; Molnar *et al.*, 1996; Schwecke *et al.*, 1995) and the non-ribosomal peptides (Miao *et al.*, 2002; Van Wageningen *et al.*, 1998).

Secondary metabolite biosynthesis can be manipulated (Hopwood *et al.*, 1985; Hutchinson, 1999; Katz & McDaniel, 1999; Kennedy & Hutchinson, 1999; McDaniel *et al.*, 1999). The ability to manipulate the biosynthetic genes for a given metabolite has diverse potential applications including the improvement of product yield in the producing organism, modification of the compound for improved activity or pharmacological properties, or the creation of completely novel hybrid compounds by combination of genes from diverse biosynthetic pathways, a concept known as combinatorial biosynthesis (Hopwood *et al.*, 1985; Hutchinson, 1999). Recently, entire genomes of the antibiotic producing *Streptomyces coelicolor* (Bentley *et al.*, 2002) and *Streptomyces avermitilis* (Ikeda *et al.*, 2003) have been sequenced, which will facilitate further studies.

Polyketides (Table 2.2) are a large, structurally diverse group of molecules that are produced by polyketide synthase (PKS) enzyme complexes. These enzymes assemble polyketide compounds by the sequential decarboxylative condensation of acyl-coenzyme A monomers, in a manner resembling the steps of fatty acid biosynthesis (Gokhale & Tuteja, 2001; Kennedy & Hutchinson, 1999; Liou & Khosla, 2003; Pfeifer & Khosla, 2001). There are two major types of PKS gene clusters. Modular type I PKSs consist of

large, multifunctional enzymes that contain multiple active sites (domains) organized into sequentially arranged groups (modules). Many important antibiotics and other secondary metabolites from both bacteria and fungi are synthesized by this type of enzyme system (Table 2.2). A type I PKS module typically contains acyltransferase (AT), ketosynthase (KS), and acyl carrier protein (ACP) domains, and may also contain ketoreductase (KR), dehydratase (DH), enoyl reductase (ER), and thioesterase (TE) domains (Figure 2.2). The AT transfers an acyl CoA precursor onto an ACP domain, after which the KS catalyzes the condensation of this precursor to a second acyl CoA extender unit, thereby extending the molecule by two or more carbons. This process is repeated as the growing molecule is transferred to subsequent modules of the PKS. In addition to these core catalytic domains, PKS modules may have other functional sites, which modify the chain during elongation by β-carbonyl reduction. A KR domain will catalyze the reduction of one of the β-ketones in the polyketide chain to a hydroxyl. A DH domain will further reduce the hydroxyl group to an olefin, and an ER domain will perform the final reduction to a methylene (Gokhale & Tuteja, 2001; Liou & Khosla, 2003). A TE domain will ultimately release the fully extended chain from the enzyme, and in many cases will also catalyze the cyclization of the chain into a macrolide ring structure (Kennedy & Hutchinson, 1999; Pfeifer & Khosla, 2001). In addition to the modular synthase components, downstream tailoring enzymes are often clustered with the PKS biosynthetic machinery. These enzymes serve to modify the polyketide core in ways that are often essential to the biological activity of the molecule, such as hydroxylation and glycosylation (Walsh, 2004).

As a result of the linear modularity of this type of PKS, the number of modules in the pathway determines the length of the final polyketide chain. One example of a modular type I PKS is 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythrea*, which synthesizes the aglycone core of erythromycin (Figure 2.2) (Pfeifer & Khosla, 2001). DEBS consists of three multifunctional proteins, each containing two modules comprised of multiple catalytic domains. The erythromycin precursor 6-deoxyerythronolide B (6-dEB) is assembled from one propionyl-CoA and six methylmalonyl-CoA precursors. The initial AT of module 1, known as the loading domain, binds the propionyl-CoA starter unit and loads it onto the ACP. Following this, the first KS catalyzes the condensation of the starter unit to a methylmalonyl-CoA extender unit. Each module contains the AT, ACP and KS catalytic domains necessary for one methylmalonyl-CoA extension, and some contain the KR, DH and ER domains necessary for reduction of the carbonyl (Figure 2.2). Biosynthesis ends with release and cyclization of the chain by a TE domain, resulting in the 14-carbon 6-dEB macrolide ring, which is subsequently glycosylated to form erythromycin A (Cortes *et al.*, 1990; Donadio *et al.*, 1991; Hutchinson, 1999).

Aromatic type II PKSs differ from modular type I PKSs in that they consist primarily of monofunctional enzymes which act in an iterative manner during polyketide synthesis. The term "minimal PKS" refers to the group of enzymes responsible for assembly of the polyketide chain, which are common to all type II PKS pathways (Figures 2.3 and 2.4, Table 2.4). The enzymes comprising the minimal PKS are as follows: a malonyl-CoA:ACP acyltransferase (MAT), a β-ketoacyl synthase (KS_α), a chain-length factor

(KS_β), and an acyl carrier protein (ACP) (Hutchinson, 1999; Keatinge-Clay *et al.*, 2004; Rawlings, 1999; Xiang *et al.*, 2004). Like the type I PKSs, extension of the polyketide chain results from condensation of acyl-CoA precursors, typically malonyl-CoA, through the cooperation of the MAT, ACP, KS_α, and KS_β. However, because type II PKS enzymes act iteratively, the length of the polyketide chain is not determined by the number of enzymes present. Instead, the length of the mature polyketide chain appears to be a function of the KS enzymes (Hutchinson, 1999; McDaniel *et al.*, 1993; Shen *et al.*, 1995). In addition to the minimal PKS, type II PKSs often contain additional enzymes including ketoreductases (KR), cyclases (CYC) and aromatases (AR) which modify the growing polyketide chain to achieve the desired aromatic structure (Xiang *et al.*, 2004; Yu *et al.*, 1998). Overall, the synthesis of type II polyketides is not well understood, and aspects of the biosynthesis such as starter unit selection and chain length determination are still controversial (Keatinge-Clay *et al.*, 2004).

The first type II PKS gene cluster to be sequenced and characterized was the actinorhodin (act) PKS from *S. coelicolor* (Hopwood *et al.*, 1985; Malpartida & Hopwood, 1984). Many commercially and medically important molecules are members of this class of compounds (Table 2.2), including the antibiotics tetracycline and doxorubicin (Hutchinson, 1999). The biosynthesis of doxorubicin is a typical example of iterative type II PKS machinery (Figure 2.3) (Pfeifer & Khosla, 2001). Doxorubicin is assembled from one propionyl-CoA starter unit plus nine malonyl-CoA extender units by the doxorubicin polyketide synthase (Dps) enzyme complex. The assembly is initiated by the condensation of the propionyl-CoA unit to a malonyl-CoA by the first set of enzymes

in the complex. These enzymes are typical type II PKS proteins, a KS_α (DpsC), a KS_β or chain-length factor (DpsD), and an ACP (DpsG). After this first diketide is formed, a second set of genes also consisting of a KS_α (DpsA), a KS_β (DpsB) and an ACP (Dps G) act together in an iterative manner, extending the chain by successive addition of eight more malonyl-CoA units. The result is a linear decaketide, which is subsequently reduced by a KR (DpsE) and ultimately cyclized by two CYCs (DpsF and DpsH) to form the aromatized core structure. Modification of the core by tailoring enzymes results in the final elaborated product (Bao *et al.*, 1999a, b).

The "minimal PKS" is similar among all type II polyketides. Diversity among type II PKSs is therefore determined by such factors as the length of chain extension, the positions at which reductions occur along the polyketide backbone, and the presence of various tailoring enzymes (Piel *et al.*, 2000a). An interesting example of unique mechanisms altering the final structure of a type II PKS product is that of the enterocin biosynthetic gene cluster (Figure 2.4) (Piel *et al.*, 2000a, 2000b; Xiang *et al.*, 2002; Xiang & Moore, 2002, 2003; Xiang *et al.*, 2004). The bacteriostat enterocin is produced by the marine bacterium *Streptomyces maritimus.* In addition to having an unusual benzoyl-CoA starter unit for the polyketide backbone, the unique core structure of this molecule is neither planar nor polyaromatic, both typical characteristics of type II polyketides. The enzymes comprising the enterocin biosynthetic pathway are shown in Table 2.4 (Piel *et al.*, 2000a). The formation of the unusually branched, rearranged enterocin polyketide appears to be derived from the action of the novel oxygenase EncM (Xiang *et al.*, 2004). The initial steps in the biosynthesis of enterocin follow the classic type II PKS paradigm,

and the "minimal PKS" genes involved (EncA, EncB, and EncC) are highly similar to those of the actinorhodin pathway (Piel *et al.*, 2000a). These enzymes construct a linear polyketide chain from the condensation of the benzoyl-CoA starter unit and seven malonyl-CoA units. EncM then catalyzes an unusual oxidative cleavage of the polyketide chain, whereby a skeletal rearrangement occurs in which several new rings and chiral centers are generated. This gene cluster is remarkable in that it does not contain typical type II PKS cyclases or aromatases. It appears instead that the unusually rearranged intermediates of enterocin biosynthesis are oriented such that the cyclizations occur spontaneously (Xiang *et al.*, 2004). The biosynthetic pathway allows for multiple paths of biosynthesis *via* alternate modifications of intermediates, which can be observed in the structurally diverse array of intermediates and shunt products produced in fermentation. For example, the deoxyenterocins result from incomplete elaboration of the molecule by tailoring enzymes, and the wailupemycins are derived from an early intermediate which spontaneously decarboxylates prior to cyclization (Piel *et al.*, 2000a).

Peptides represent another group of commercially important secondary metabolites (Table 2.3). Some of these peptides are produced by microbial modular synthetases (Marahiel *et al.*, 1997), and include the β-lactams, such as penicillin, the glycopeptides, such as vancomycin, and the lipopeptides, such as daptomycin. The biosynthesis of these peptides occurs non-ribosomally and is catalyzed by a multifunctional enzyme complex, the non-ribosomal peptide synthase (NRPS). The NRPS is arranged in a linear, modular format similar to the type I PKS, in which synthase subunits contain multiple modules each responsible for a single step of elongation, as catalyzed by the catalytic domains

comprising the module (Kleinkauf & von Dohren, 1996). The peptide is assembled from activated (adenylated) amino acid precursors, much in the same way that type I polyketides are assembled from acyl-CoA monomers, and the non-ribosomal nature of biosynthesis allows for the inclusion of unusual non-protein amino acids and D-isomers. The catalytic domains of a typical NRPS are functionally similar to those of a type I PKS. An adenylation (A) domain selects a specific amino acid precursor, activates it by acylation, and transfers the newly formed aminoacyl to the peptidyl carrier protein (PCP) domain (also referred to as a thiolation or T domain) (Walsh, 2003). The condensation (C) domain then catalyzes peptide bond formation between the PCP-bound aminoacyl group and the amino acid bound on the next (downstream) PCP module (Kennedy & Hutchinson, 1999). Elongation occurs stepwise in a linear fashion, with the number of modules dictating the length of the peptide backbone. The module responsible for chain termination includes a thioesterase (TE) domain similar to those found in type I PKSs, which will catalyze the release of the fully extended peptide chain from the synthase, and in many cases will also catalyze macrocyclization of the chain (Walsh, 2004).

Modification of the peptide backbone can result from the presence of additional domains within synthetase modules or tailoring enzymes clustered with the pathway. Epimerization (E) domains catalyze the conversion of L-amino acid precursors to their D-isomers (Kleinkauf & von Dohren, 1996; Walsh, 2003). Other domains catalyze N-methylation, acylation or cyclization of individual amino acids or portions of the chain (Kennedy & Hutchinson, 1999; Kleinkauf & von Dohren, 1997; von Dohren & Kleinkauf, 1997). And, as with polyketides, tailoring enzymes are responsible for final

steps in biosynthesis such as glycosylation or hydroxylation, which are often necessary for the biological activity of the final product (Walsh, 2004).

The classic example of an NRPS biosynthetic pathway is that of the ACV (aminoadipyl-cysteinyl-valine) synthetase (Figure 2.5) (Cane *et al.*, 1998). This pathway synthesizes the tripeptide ACV, which is the common precursor to penicillin, cephalosporin and cephamycin biosynthesis in actinomycetes and fungi. ACV synthetase is comprised of 10 domains, which catalyze the condensation of three amino acids. The initiation module selects and activates an L-α-aminoadipate, after which two subsequent modules catalyze the addition of L-cysteine and D-valine to form the tripeptide. The initiation module and the first elongation module contain the standard three domains (A, PCP, C) for amino acid selection and condensation. The module responsible for valine addition also contains an E domain, which catalyzes the epimerization of L-valine to D-valine, and a TE domain which terminates elongation (Paradkar *et al.*, 1997). The end product is the tripeptide ACV, which is converted to isopenicillin N and then further modified to yield the various β-lactam antibiotics mentioned above.

Advances in our understanding of the biosynthetic pathways for important secondary metabolites, coupled with new technologies for analyzing and manipulating gene clusters, have both revolutionized and energized the field of natural products research. A highly desirable application of these recent breakthroughs is the creation of new molecular entities through the manipulation of biosynthesis. Production of novel, "unnatural" natural products, with potential improvements in activity, potency, or other

pharmacological properties compared to naturally occurring compounds, has been studied for decades. Prior to the molecular manipulation of biosynthetic genes, this was primarily achieved through precursor-directed biosynthesis. By this method, modifications to the structure of a secondary metabolite are made possible by exploitation of the relaxed substrate specificity of certain enzymes in the biosynthetic pathway. Analogs of known precursors are supplied in the culture medium, and are subsequently incorporated into the growing molecule in place of the natural precursor, resulting in a modified final product. Well-known examples of this include the replacement of the pipecolate moiety of the polyketide rapamycin with proline, as well as with certain proline and pipecolate analogs (Graziani *et al.*, 2003; Khaw *et al.*, 1998; Kojima & Demain, 1998; Ritacco *et al.*, 2005). The important role of the pipecolate moiety in the binding of rapamycin to its target, FKBP, has made this region of the molecule a desirable target for modification. Other such modifications to the rapamycin molecule have been achieved by the feeding of unusual starter units, which replace the native shikimic acid derived starter (Lowden *et al.*, 2004).

Direct molecular manipulation of biosynthetic genes makes it possible to target specific portions of a molecule for modification, and to create a series of novel analogs of the compound by alteration of the enzymes that build the molecular skeleton or the tailoring enzymes that decorate the core in order to yield the final product. Perhaps the simplest form of this type of manipulation is the direct knockout of a biosynthetic gene (or genes) responsible for the synthesis or incorporation of specific molecular precursors. Elimination of the supply of a certain precursor has been shown to facilitate precursor-

directed biosynthesis, by eliminating competition with the natural (desired) substrate. For example, in the rapamycin biosynthetic gene cluster, the gene *rapL* encodes a lysine cyclodeaminase that catalyzes the conversion of lysine to pipecolate (Aparicio *et al.*, 1996; Molnar *et al.*, 1996; Paiva *et al.*, 1993; Schwecke *et al.*, 1995). Deletion of this gene from the pathway dramatically reduces the availability of pipecolate for rapamycin biosynthesis, and thereby increases the incorporation of proline and certain proline analogs into rapamycin molecule in place of pipecolate (Khaw *et al.*, 1998). This correlates with other studies in which similar results were achieved by direct inhibition of the lysine cyclodeaminase enzyme, rather than genetic knockout (Graziani *et al.*, 2003; Kojima & Demain, 1998; Ritacco *et al.*, 2005).

The term combinatorial biosynthesis has been coined to describe the creation of novel hybrid compounds by combining genes from different biosynthetic gene clusters, resulting in the biosynthesis of new end products not previously seen in nature (Hopwood *et al.*, 1985; Hutchinson, 1999). This approach holds great promise in the search for novel and improved antibacterial compounds effective against highly resistant bacterial pathogens, and will provide libraries of novel structural analogs of bioactive secondary metabolites for high throughput screening against a range of pharmaceutical targets. Combinatorial biosynthesis was first demonstrated by Hopwood *et al.* (Hopwood *et al.*, 1985) who introduced some of the genes from the actinorhodin (ACT) type II PKS biosynthetic pathway of *S. coelicolor* into other streptomycetes known to produce the structurally related type II polyketides granaticin and medermycin. These experiments resulted in the production of novel hybrid molecules, clearly derived from the interaction

of genes from different pathways, and from different producing organisms. More recently, researchers at Kosan Biosciences (Katz & McDaniel, 1999; McDaniel *et al.*, 1999) made a large number of modifications to the erythromycin PKS of *S. erythrea*, resulting in greater than 50 novel structural variations. The erythromycin pathway has become a model system for studying these types of manipulations, and mutations have been demonstrated in most of the catalytic domains of the PKS, yielding modified structures (Walsh, 2003).

Another application of our increasing understanding of the genetics of secondary metabolite biosynthesis is the ability to search for specific types of biosynthetic genes within the genomes of microorganisms or in environmental samples. The genomic DNA of an organism can be screened for the presence of specific biosynthetic genes by the use of radiolabeled probes or PCR primers based on known sequences from related pathways. This technique has been used to prescreen microorganisms to identify those with the biosynthetic potential to produce secondary metabolites of interest, and has also been used for the cloning and characterization of biosynthetic pathways (Anderson *et al.*, 2002; Ayuso-Sacido & Genilloud, 2005; Decker *et al.*, 1996; Metsa-Ketela *et al.*, 1999; Piel *et al.*, 2000a; Ritacco *et al.*, 2003; Sigmund *et al.*, 2003; Sosio *et al.*, 2000). The same method has also been used to directly screen environmental samples, such as soil, for biosynthetic pathways in order to assess the quantity and diversity of these pathways in the sample prior to actual isolation of microorganisms. This is accomplished either by PCR amplification of genes from DNA isolated directly from an environmental sample, or by screening “metagenomic” libraries generated by the shotgun cloning of random

DNA sequences isolated from the environment (Courtois *et al.*, 2003; Handelsman *et al.*, 2002; Handelsman, 2004; Rondon *et al.*, 2000; Schloss & Handelsman, 2003; Wawrik *et al.*, 2005). These types of approaches provide a glimpse of the metabolic potential present in the environment independent of actual microbial cultivation, thereby allowing investigators to focus on the samples with the greatest metabolic diversity, and making possible the identification and expression of novel metabolites from previously "unculturable" microorganisms.

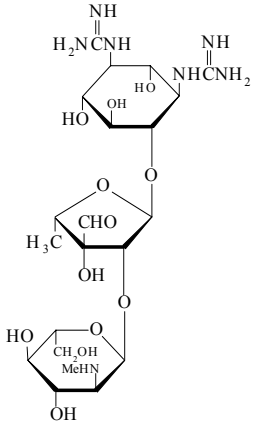
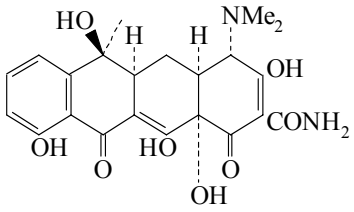
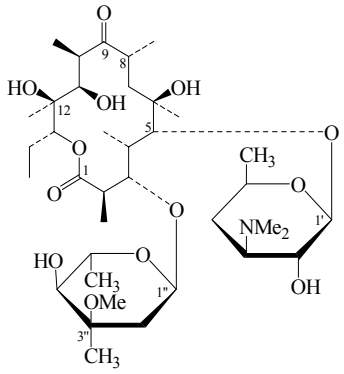
Name	Type	Structure	Producer	Activity
Streptomycin	Aminoglycoside		*Streptomyces griseus*	Antibacterial
Tetracycline	Type II polyketide		*Streptomyces aureofaciens*	Antibacterial
Erythromycin	Type I polyketide; glycosylated macrolide		*Saccharopolyspora erythrea*	Antibacterial

Table 2.1. Examples of commercially important antibiotics from actinomycetes. Structures are from The Dictionary of Natural Products on CD-ROM, Chapman and Hall/CRC (Buckingham, 2004).

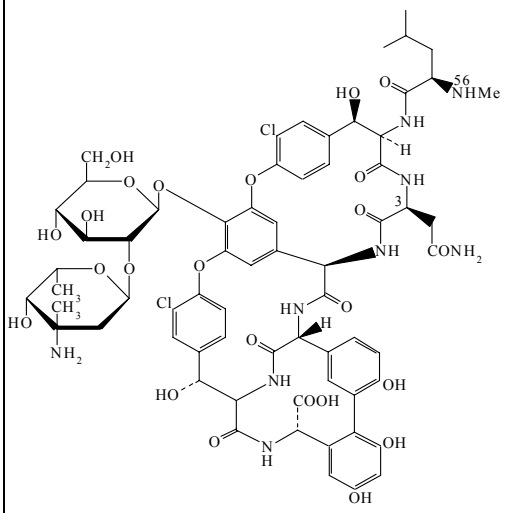
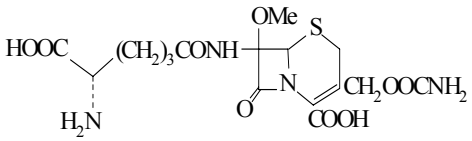
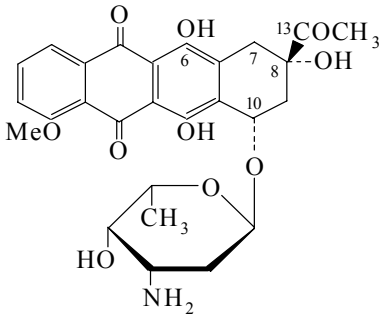
Name	Type	Structure	Producer	Activity
Vancomycin	Glycopeptide		*Amycolatopsis orientalis*	Antibacterial
Cephamycin C	β-lactam		*Streptomyces clavuligerus*	Antibacterial
Daunorubicin	Type II polyketide; anthracycline-like		*Streptomyces peucetius*	Anticancer

Table 2.1 continued. Examples of commercially important antibiotics from actinomycetes**.** (Buckingham, 2004).

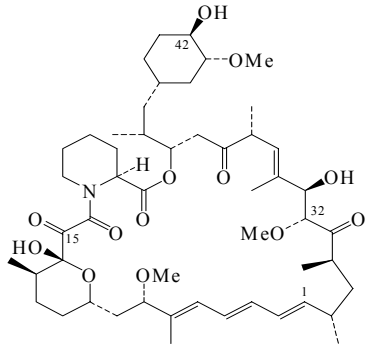
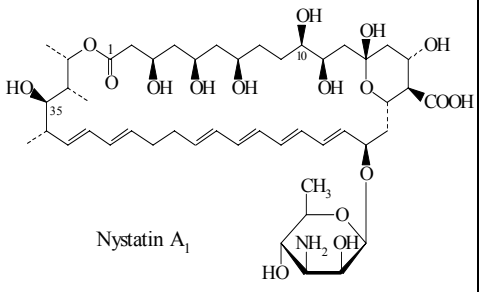
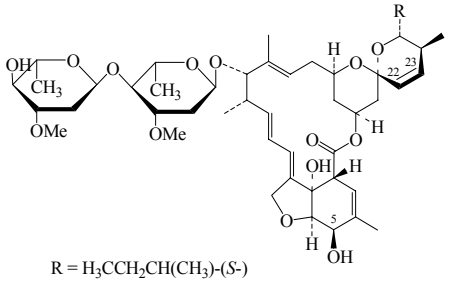
Name	Type	Structure	Producer	Activity
Rapamycin	Type I polyketide / non-ribosomal peptide hybrid		*Streptomyces hygroscopicus*	Immunosuppressant; anticancer; antifungal
Nystatin	Polyene	 Nystatin A_1	*Streptomyces noursei*	Antifungal
Avermectin B	Type I polyketide; glycosylated macrolide	 R = $H_3CCH_2CH(CH_3)$-(*S*-)	*Streptomyces avermitilis*	Antiparasitic

Table 2.1 continued. Examples of commercially important antibiotics from actinomycetes (Buckingham, 2004).

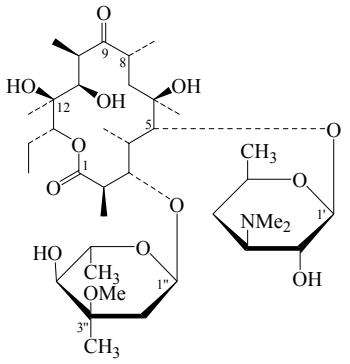
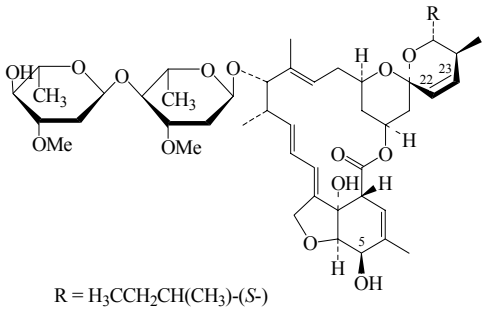
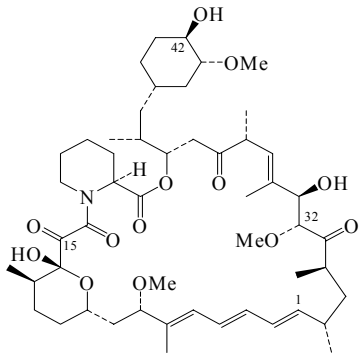
Name	Polyketide type	Structure	Activity / Function
Erythromycin	Type I polyketide; glycosylated macrolide		Antibacterial
Avermectin B	Type I polyketide; glycosylated macrolide	 R = $H_3CCH_2CH(CH_3)$-(*S*-)	Antiparasitic
Rapamycin	Type I polyketide/ non-ribosomal peptide hybrid; macrolide		Immunosuppressant; anticancer; antifungal

Table 2.2. Structural diversity of polyketides. Structures are from The Dictionary of Natural Products on CD-ROM, Chapman and Hall/CRC (Buckingham, 2004).

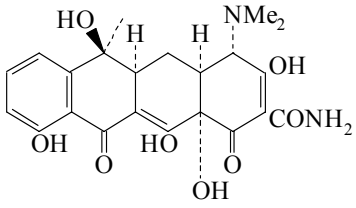
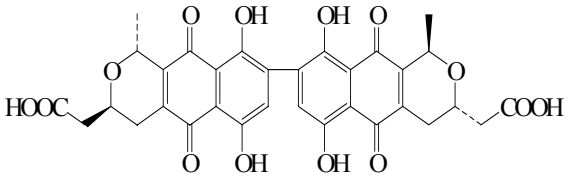
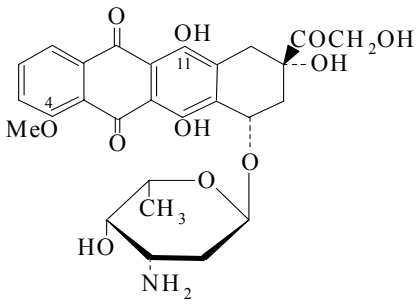
Name	Polyketide type	Structure	Activity / Function
Tetracycline	Aromatic type II polyketide		Antibacterial
Actinorhodin	Aromatic type II polyketide		Pigment; pH indicator
Doxorubicin	Aromatic type II polyketide; anthracycline-like		Anticancer

Table 2.2 continued. Structural diversity of polyketides (Buckingham, 2004).

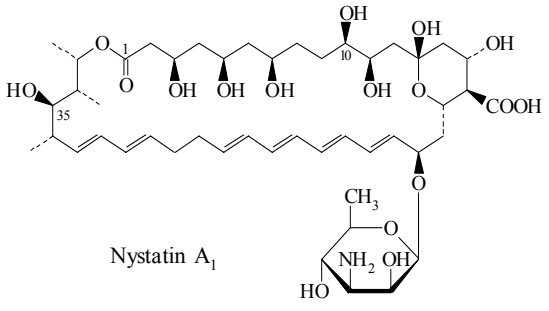
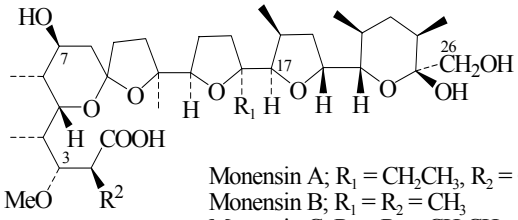
Name	Polyketide type	Structure	Activity / Function
Nystatin	Polyene	 Nystatin A_1	Antifungal
Monensin	Polyether	 Monensin A; $R_1 = CH_2CH_3$, $R_2 = CH_3$ Monensin B; $R_1 = R_2 = CH_3$ Monensin C; $R_1 = R_2 = CH_2CH_3$ 2-Demethylmonensin A; $R_1 = CH_2CH_3$, $R_2 = H$ 2-Demethylmonensin B; $R_1 = CH_3$, $R_2 = H$	Antiparasitic

Table 2.2 continued. Structural diversity of polyketides (Buckingham, 2004).

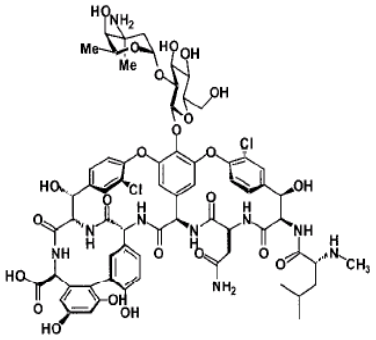
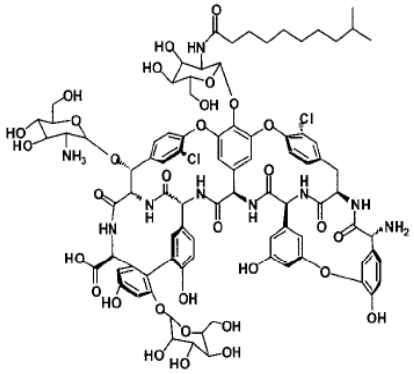
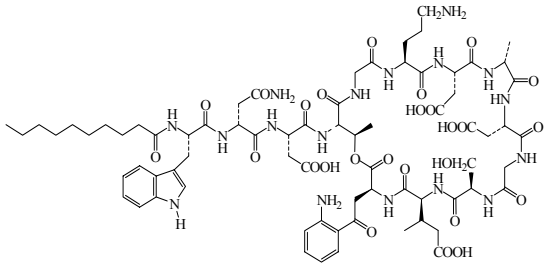
Name	Biosynthesis	Peptide type	Structure	Activity / Function
Vancomycin	Non-ribosomal peptide synthetase (NRPS)	Glycopeptide		Antibacterial
Teicoplanin A	Non-ribosomal peptide synthetase (NRPS)	Lipoglycopeptide		Antibacterial
Daptomycin	Non-ribosomal peptide synthetase (NRPS)	Lipopeptide		Antibacterial

Table 2.3. Structural diversity of peptide antibiotics (Buckingham, 2004; Hubbard & Walsh, 2003).

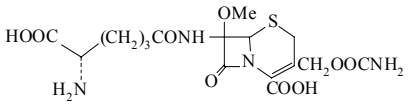
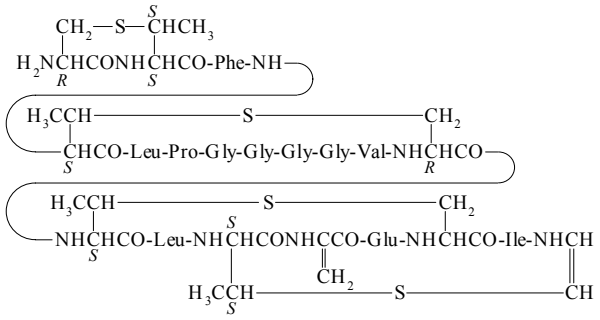
Name	Biosynthesis	Peptide type	Structure	Activity / Function
Cephamycin C	Non-ribosomal peptide synthetase (NRPS)	β-lactam		Antibacterial
Mersacidin	Ribosomal	Lantibiotic	 Proposed structure	Antibacterial

Table 2.3 continued. Structural diversity of peptide antibiotics (Buckingham, 2004).

Protein	Amino acids (no.)	Proposed function
EncA	423	KS_{α}
EncB	407	KS_{β}
EncC	82	ACP
EncD	266	KR
EncE	163	Regulatory protein
EncF	283	Transcriptional activator
EncG	163	Unknown
EncH	535	Acyl-CoA ligase
EncI	258	Enoyl-CoA hydratase
EncJ	400	β-Oxoacyl-CoA thiolase
EncK	241	Methyltransferase (MT)
EncL	340	AT
EncM	464	FAD-dependent oxygenase
EncN	522	Aryl-CoA ligase
EncO	130	Unknown
EncP	523	PAL
EncQ	81	Ferredoxin
EncR	401	Cytochrome P-450 hydroxylase
EncS	215	Regulatory protein
EncT	482	Efflux protein

Table 2.4. The enzymes of the enterocin (Enc) biosynthetic pathway.

KS= ketosynthase, ACP= acyl carrier protein, KR= ketoreductase, MT= methyltransferase, AT= acyltransferase, PAL= phenylalanine ammonia lyase.

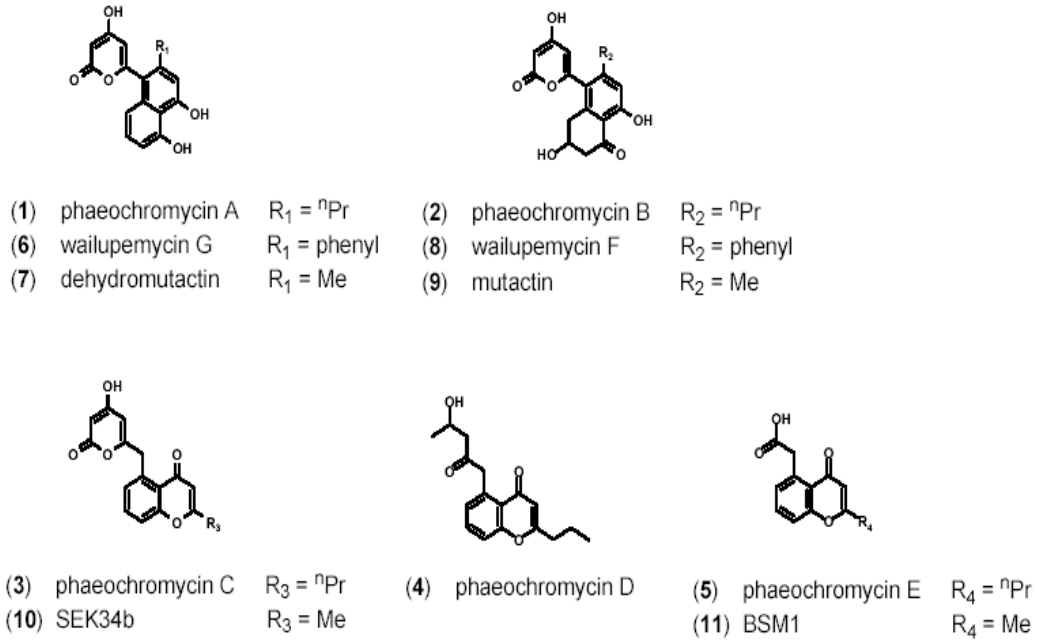


Figure 2.1. Structures of phaeochromycins A-E (1 to 5) and related metabolites. The phaeochromycins are analogous to shunt metabolites from the actinorhodin and enterocin biosynthetic pathways, namely wailupemycin G (6), dehydromutactin (7), wailupemycin F (8), mutactin (9), SEK34b (10), and BSM1 (11). The phaeochromycins differ from these analogous metabolites by the presence of a C4 starter unit, presumably derived from butyrate. Me = methyl, ^{n}Pr = n-propyl (Graziani *et al.*, 2005; Hopwood, 1997; Kalaitzis & Moore, 2004; McDaniel *et al.*, 1994a; McDaniel *et al.*, 1994b; Xiang *et al.*, 2002).

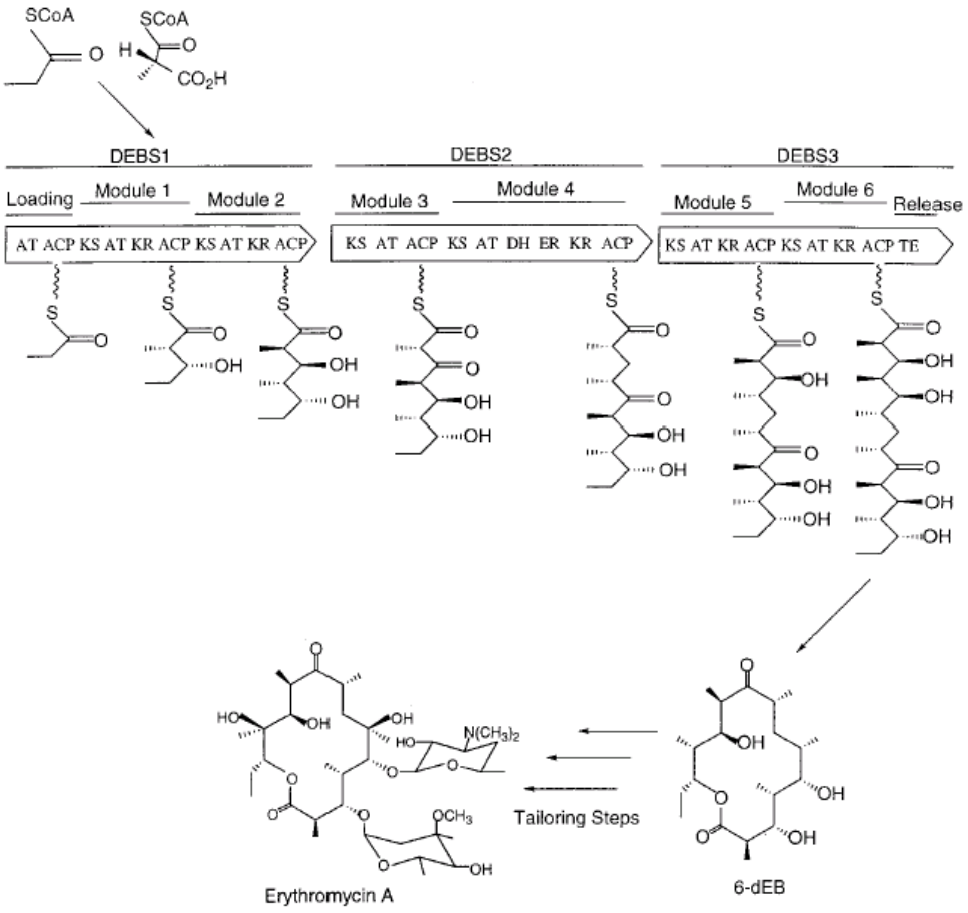


Figure 2.2. The erythromycin biosynthetic pathway (Pfeifer & Khosla, 2001). The 6-deoxyerythronolide B synthase (6dEBS) consists of three multifunctional proteins (DEBS1, DEBS2, DEBS3) each containing two modules comprised of multiple catalytic domains. This enzyme complex synthesizes the erythromycin precursor 6-deoxyerythronolide B (6dEB) from one propionyl-CoA starter unit and 6 methylmalonyl-CoA molecules. AT= acyltransferase, ACP= acyl carrier protein, KS= ketosynthase, KR=ketoreductase, DH= dehydratase, ER= enoyl reductase, TE= thioesterase.

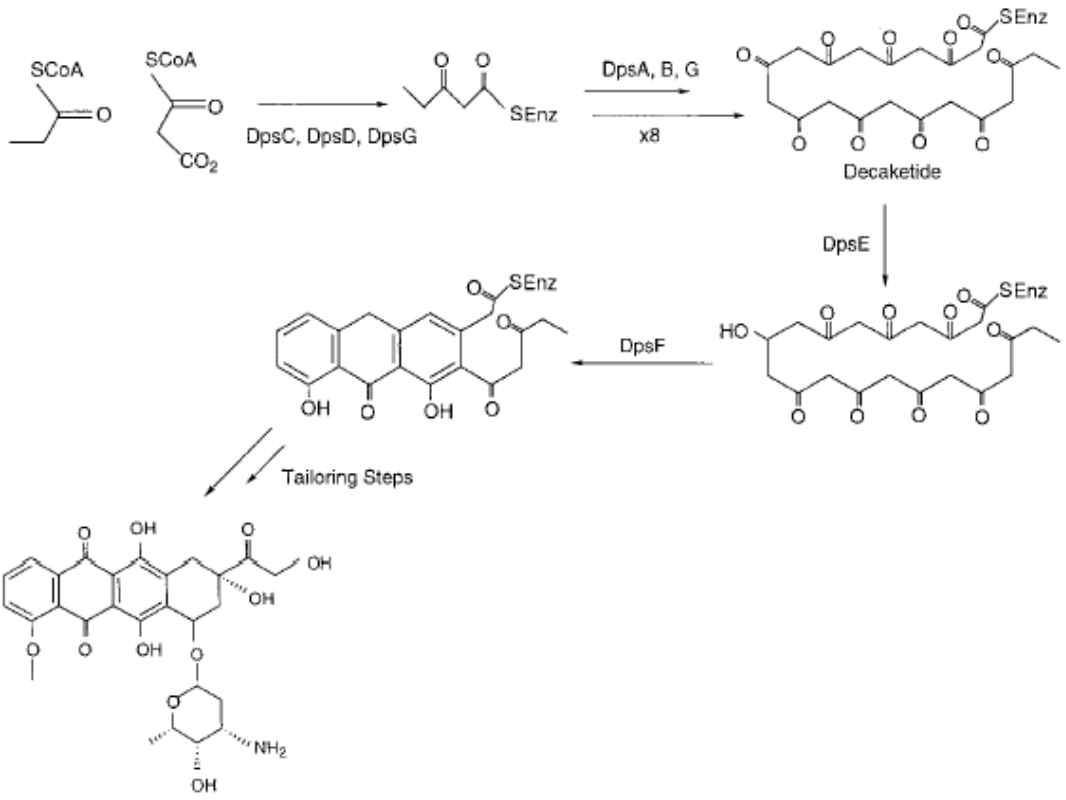


Figure 2.3. The biosynthesis of doxorubicin (Pfeifer & Khosla, 2001). Doxorubicin is assembled from a propionyl-CoA starter unit plus nine malonyl-CoA extender units by the doxorubicin polyketide synthase (Dps) enzyme complex.

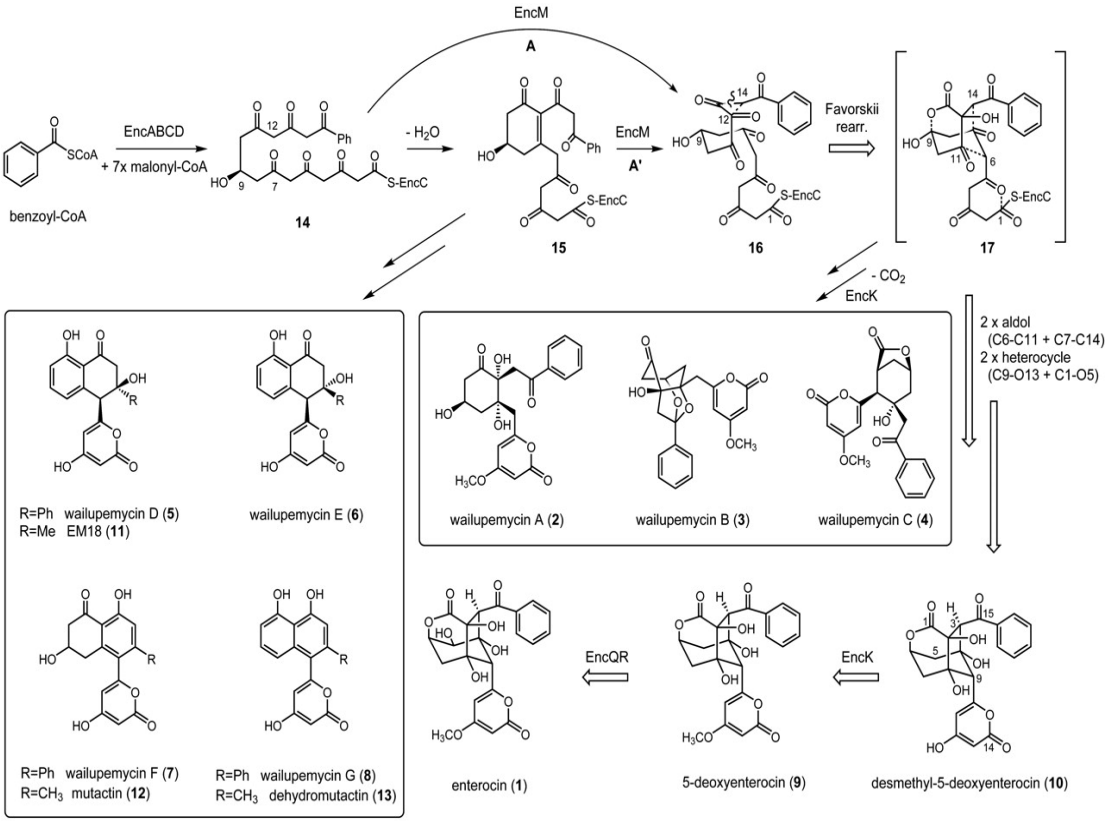


Figure 2.4. Biosynthesis of enterocin and the wailupemycins (Xiang *et al.*, 2004). The biosynthetic complex is a type II PKS, and is comprised of the enzymes EncA-T. The proposed function of each of these enzymes is detailed in Table 2.4.

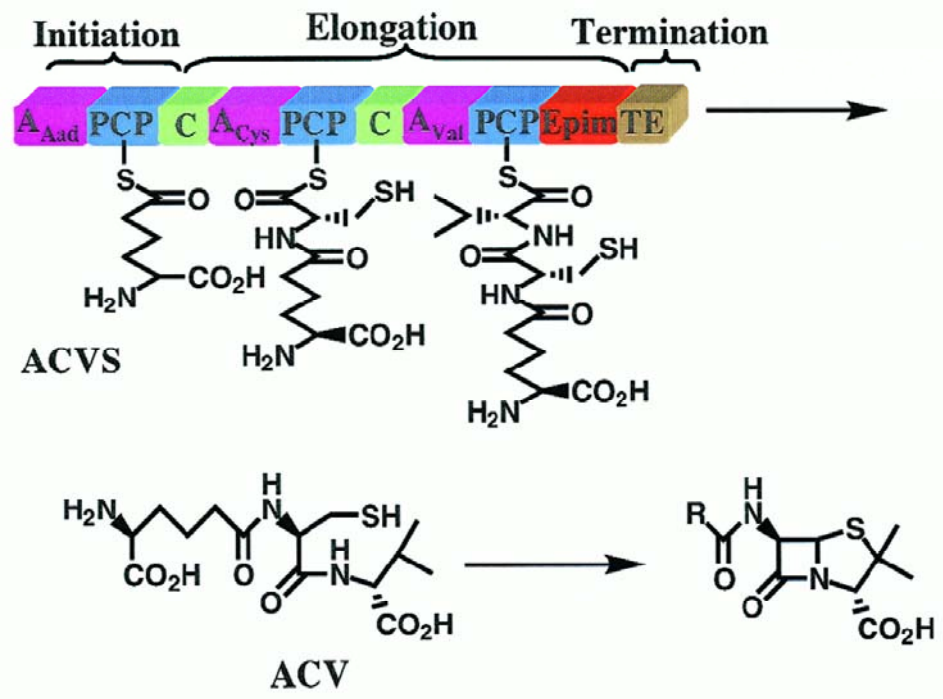


Figure 2.5. ACV synthetase (Cane *et al.*, 1998) and the synthesis of the tripeptide ACV (aminoadipyl-cysteinyl-valine).

ACV is a precursor to penicillin, cephalosporin and cephamycin biosynthesis. A= adenylation domain, PCP= peptidyl carrier protein domain, C= condensation domain, Epim= epimerization domain, TE= thioesterase domain, Aad= aminoadipate, Cys= cysteine, Val= valine.

References

Anderson, A. S., Clark, D. J., Gibbons, P. H. & Sigmund, J. M. (2002). The detection of diverse aminoglycoside phosphotransferases within natural populations of actinomycetes. *J Ind Microbiol Biotechnol* **29**, 60-69.

Anderson, A. S. & Wellington, E. M. H. (2001). The taxonomy of *Streptomyces* and related genera. *Int J Syst Evol Microbiol* **51**, 797-814.

Anzai, Y., Okuda, T. & Watanabe, J. (1997). Application of the random amplified polymorphic DNA using the polymerase chain reaction for efficient elimination of duplicate strains in microbial screening. II. Actinomycetes. *J Antibiot* **50**, C-1.

Aparicio, J. F., Molnar, I., Schwecke, T., Koenig, A., Haydock, S. F., Khaw, L. E., Staunton, J. & Leadlay, P. F. (1996). Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymic domains in the modular polyketide synthase. *Gene* **169**, 9-16.

Ayuso-Sacido, A. & Genilloud, O. (2005). New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb Ecol* **49**, 10-24.

Baltz, R. H. (1997). Lipopeptide antibiotics produced by *Streptomyces roseosporus* and *Streptomyces fradiae*. In *Drugs and the Pharmaceutical Sciences*, pp. 415-435. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Baltz, R. H. (2006). Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J Ind Microbiol Biotechnol* **33**, 507-513.

Bao, W., Sheldon, P. J. & Hutchinson, C. R. (1999a). Purification and properties of the *Streptomyces peucetius* DpsC beta-ketoacyl:acyl carrier protein synthase III that specifies the propionate-starter unit for type II polyketide biosynthesis. *Biochemistry* **38**, 9752-9757:.

Bao, W., Sheldon, P. J., Wendt-Pienkowski, E. & Hutchinson, C. R. (1999b). The *Streptomyces peucetius* dpsC gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide. *J Bacteriol* **181**, 4690-4695.

Bassler, B. L. & Losick, R. (2006). Bacterially speaking. *Cell* **125**, 237-246.

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J. & Hopwood, D. A. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.

Bernan, V. S., Montenegro, D. A., Goodman, J. J., Alluri, M. R., Carter, G. T., Abbanat, D. R., Pearce, C. J., Maiese, W. M. & Greenstein, M. (1994). Martinomycin, a new polyether antibiotic produced by *Streptomyces salvialis*. I. Taxonomy, fermentation and biological activity. *J Antibiot* **47**, 1434-1441.

Buckingham, J. (2004). *Dictionary of Natural Products on CD-ROM*.Chapman & Hall/CRC, Boca Raton, FL.

Cane, D. E., Walsh, C. T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63-68.

Challis, G. L. & Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci USA* **100 (Suppl 2)**, 14555-14561.

Chater, K. F. & Bibb, M. J. (1997). Regulation of bacterial antibiotic production. In *Biotechnology (2nd Edition)*, pp. 57-105. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Coisne, S., Bechet, M. & Blondeau, R. (1999). Actinorhodin production by *Streptomyces coelicolor* A3(2) in iron-restricted media. *Lett Appl Microbiol* **28**, 199-202.

Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**, 176-178.

Courtois, S., Cappellano, C. M., Ball, M., Francou, F.-X., Normand, P., Helynck, G., Martinez, A., Kolvek, S. J., Hopke, J., Osburne, M. S., August, P. R., Nalin, R.,

Guerineau, M., Jeannin, P., Simonet, P. & Pernodet, J.-L. (2003). Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* **69**, 49-55.

Dahod, S. K. (1999). Raw materials selection and medium development for industrial fermentation processes. In *Manual of Industrial Microbiology and Biotechnology, Second Edition*, pp. 213-220. Ed. A. L. Demain & J. E. Davies. ASM Press, Washington, D.C.

Decker, H., Gaisser, S., Pelzer, S., Schneider, P., Westrich, L., Wohlleben, W. & Bechthold, A. (1996). A general approach for cloning and characterizing dNDP-glucose dehydratase genes from actinomycetes. *FEMS Microbiol Lett* **141**, 195-201.

Demain, A. L. & Vaishnav, P. (2004). Secondary metabolism in microbes and its control by phosphate and metals. *SIM News* **54**, 104-113.

Distler, J., Mansouri, K., Mayer, G., Stockmann, M. & Piepersberg, W. (1992). Streptomycin biosynthesis and its regulation in streptomycetes. *Gene* **115**, 105-111.

Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.

Fernandez-Moreno, M. A., Martinez, E., Boto, L., Hopwood, D. A. & Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J Biol Chem* **267**, 19278-19290.

Gokhale, R. S. & Tuteja, D. (2001). Biochemistry of polyketide synthases. In *Biotechnology (2nd Edition)*, pp. 341-372. Ed. H. Kleinkauf & H. Von Doehren.

Goodfellow, M. (1989). The Actinomycetes I. Suprageneric classification of actinomycetes. In *Bergey's Manual of Systematic Bacteriology*, pp. 2333-2339. Ed. S. T. Williams, M. E. Sharpe & J. G. Holt. Williams & Wilkins, Baltimore, MD.

Goodfellow, M., Ferguson, E. V. & Sanglier, J. J. (1992). Numerical classification and identification of *Streptomyces* species--a review. *Gene* **115**, 225-233.

Graziani, E. I., Ritacco, F. V., Bernan, V. S. & Telliez, J. B. (2005). Phaeochromycins A-E, anti-inflammatory polyketides isolated from the soil actinomycete *Streptomyces phaeochromogenes* LL-P018. *J Nat Prod* **68**, 1262-1265.

Graziani, E. I., Ritacco, F. V., Summers, M. Y., Zabriskie, T. M., Yu, K., Bernan, V. S., Greenstein, M. & Carter, G. T. (2003). Novel sulfur-containing rapamycin analogs prepared by precursor-directed biosynthesis. *Org Lett* **5**, 2385-2388.

Handelsman, J. (2004). Soils - the metagenomics approach. In *Microbial Diversity and Bioprospecting*, pp. 109-119. Ed. A. T. Bull. ASM Press, Herndon, VA.

Handelsman, J., Liles, M., Mann, D., Riesenfeld, C. & Goodman, R. M. (2002). Cloning the metagenome: culture-independent access to the diversity and functions of the uncultivated microbial world. *Methods in Microbiology* **33**, 241-255.

Hansen, J. N. (1997). Nisin and related antimicrobial peptides. In *Drugs and the Pharmaceutical Sciences*, pp. 437-470. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Hatano, K., Nishii, T. & Kasai, H. (2003). Taxonomic re-evaluation of whorl-forming *Streptomyces* (formerly *Streptoverticillium*) species by using phenotypes, DNA-DNA hybridization and sequences of *gyrB*, and proposal of *Streptomyces luteireticuli* (ex Katoh and Arai 1957) corrig., sp. nov., nom. rev. *Int J Syst Evol Microbiol* **53**, 1519-1529.

Hopwood, D. A. (1997). Genetic contributions to understanding polyketide synthases. *Chem Rev* **97**, 2465-2497.

Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fujii, I., Rudd, B. A. M., Floss, H. G. & Omura, S. (1985). Production of 'hybrid' antibiotics by genetic engineering. *Nature (London, United Kingdom)* **314**, 642-644.

Hubbard, B. K. & Walsh, C. T. (2003). Vancomycin assembly: nature's way. *Angew Chem* **42**, 730-765.

Hunter-Cevera, J. C. & Belt, A. (1999). Isolation of cultures. In *Manual of Industrial Microbiology and Biotechnology, Second Edition*. Ed. A. L. Demain & J. E. Davies. ASM Press, Washington, D.C.

Hutchinson, C. R. (1999). Combinatorial biosynthesis of antibiotics. In *Drug Discovery from Nature*, pp. 233-254. Eds. S. Grabley & R. Thiericke. Springer-Verlag, Berlin, Germany.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**, 526-531.

Jack, R., Gotz, F. & Jung, G. (1997). Lantibiotics. In *Biotechnology (2nd Edition)*, pp. 323-368. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Kalaitzis, J. A. & Moore, B. S. (2004). Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* **67**, 1419-1422.

Katz, L. & McDaniel, R. (1999). Novel macrolides through genetic engineering. *Med Res Rev* **19**, 543-558.

Kawabuchi, M., Hara, Y., Nihira, T. & Yamada, Y. (1997). Production of butyrolactone autoregulators by *Streptomyces coelicolor* A3(2). *FEMS Microbiol Lett* **157**, 81-85.

Keatinge-Clay, A. T., Maltby, D. A., Medzihradszky, K. F., Khosla, C. & Stroud, R. M. (2004). An antibiotic factory caught in action. *Nat Struct Mol Biol* **11**, 888-893.

Kennedy, J. & Hutchinson, C. R. (1999). Nurturing nature: engineering new antibiotics. *Nat Biotechnol* **17**, 538-539.

Khaw, L. E., Bohm, G. A., Metcalfe, S., Staunton, J. & Leadlay, P. F. (1998). Mutational biosynthesis of novel rapamycins by a strain of *Streptomyces hygroscopicus* NRRL 5491 disrupted in *rapL*, encoding a putative lysine cyclodeaminase. *J Bacteriol* **180**, 809-814.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces genetics*.The John Innes Foundation, Norwich, England.

Kieser, T. & Hopwood, D. A. (1991). Genetic manipulation of *Streptomyces*: integrating vectors and gene replacement. *Methods Enzymol* **204**, 430-458.

Kim, B. J., Kim, C. J., Chun, J., Koh, Y. H., Lee, S. H., Hyun, J. W., Cha, C. Y. & Kook, Y. H. (2004). Phylogenetic analysis of the genera *Streptomyces* and *Kitasatospora* based on partial RNA polymerase beta-subunit gene (*rpoB*) sequences. *Int J Syst Evol Microbiol* **54**, 593-598.

Kleinkauf, H. & von Dohren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur J Biochem* **236**, 335-351.

Kleinkauf, H. & von Dohren, H. (1997). Peptide antibiotics. In *Biotechnology (2nd Edition)*, pp. 277-322. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Kojima, I. & Demain, A. L. (1998). Preferential production of rapamycin versus prolylrapamycin by *Streptomyces hygroscopicus*. *J Ind Microbiol Biotechnol* **20**, 309-316.

Labeda, D. P. (1992). DNA-DNA hybridization in the systematics of Streptomyces. *Gene* **115**, 249-253.

Lanoot, B., Vancanneyt, M., Hoste, B., Vandemeulebroecke, K., Cnockaert, M. C., Dawyndt, P., Liu, Z., Huang, Y. & Swings, J. (2005). Grouping of streptomycetes using 16S-ITS RFLP fingerprinting. *Res Microbiol* **156**, 755-762.

Lechevalier, M. P. (1977). Lipids in bacterial taxonomy - a taxonomist's view. *Crit Rev Microbiol* **5**, 109-210.

Lechevalier, M. P. & Lechevalier, H. (1970). Chemical composition as a criterion in the classification of aerobic actinomycetes. *Int J Syst Bacteriol* **20**, 435-443.

Liou, G. F. & Khosla, C. (2003). Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol* **7**, 279-284.

Lowden, P. A. S., Boehm, G. A., Metcalfe, S., Staunton, J. & Leadlay, P. F. (2004). New rapamycin derivatives by precursor-directed biosynthesis. *ChemBioChem* **5**, 535-538.

Malpartida, F. & Hopwood, D. A. (1984). Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature (London)* **309**, 462-464.

Mankelow, D. P. & Neilan, B. A. (2000). Non-ribosomal peptide antibiotics. *Exp Op Therap Pat* **10**, 1583-1591.

Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem Rev* **97**, 2651-2673.

McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A. & Khosla, C. (1994a). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci USA* **91**, 11542-11546.

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1993). Engineered biosynthesis of novel polyketides. *Science* **262**, 1546-1550.

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1994b). Engineered biosynthesis of novel polyketides: actVII and actIV genes encode aromatase and cyclase enzymes, respectively. *J Am Chem Soc* **116**, 10855-10859.

McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Betlach, M. & Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc Natl Acad Sci USA* **96**, 1846-1851.

Metsa-Ketela, M., Salo, V., Halo, L., Hautala, A., Hakala, J., Mantsala, P. & Ylihonko, K. (1999). An efficient approach for screening minimal PKS genes from *Streptomyces*. *FEMS Microbiol Lett* **180**, 1-6.

Miao, V. P. W., Brian, P., Baltz, R. H. & Silva, C. J. (2002). Daptomycin biosynthetic gene cluster from *Streptomyces roseosporus*. In *PCT Int Appl*, pp. 227 pp. Wo: (Cubist Pharmaceuticals, Inc., Can.).

Miller, T. L. & Churchill, B. W. (1986). Substrates for large scale fermentations. In *Manual of Industrial Microbiology and Biotechnology, First Edition*. Ed. A. L. Demain & N. A. Solomon. ASM Press, Washington, D.C.

Minnikin, D. E., O'Donnell, A. G., Goodfellow, M., Alderson, G., Athalye, M., Schaal, A. & Parlett, J. H. (1984). An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* **2**, 233-241.

Molnar, I., Aparicio, J. F., Haydock, S. F., Khaw, L. E., Schwecke, T., Koenig, A., Staunton, J. & Leadlay, P. F. (1996). Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase. *Gene* **169**, 1-7.

Moss, C. W. & Guerrant, G. O. (1983). Separation of bacterial ubiquinones by reverse-phase high-pressure liquid chromatography. *J Clin Microbiol* **18**, 15-17.

Nagasawa, T., Ling, J. R. & Onodera, R. (1993). Chiral high-performance liquid chromatographic separation of the three stereoisomers of 2,6-diaminopimelic acid without derivatization. *J Chromatogr* **653**, 336-340.

Nicas, T. I. & Cooper, R. D. G. (1997). Vancomycin and other glycopeptides. In *Drugs and the Pharmaceutical Sciences*, pp. 363-392. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Nonomura, H. (1974). Key for the classification and identification of 458 species of the Streptomycetes included in ISP. *J Ferm Tech* **52**, 78-92.

Omura, S. (1992). *The search for bioactive compounds from microorganisms*.Springer-Verlag, Berlin, West Germany.

Paiva, N. L., Demain, A. L. & Roberts, M. F. (1993). The immediate precursor of the nitrogen-containing ring of rapamycin is free pipecolic acid. *Enzyme Microb Technol* **15**, 581-585.

Paradkar, A. S., Jensen, S. E. & Mosher, R. H. (1997). Comparative genetics and molecular biology of b-lactam biosynthesis. In *Drugs and the Pharmaceutical Sciences*, pp. 241-277. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Pfeifer, B. A. & Khosla, C. (2001). Biosynthesis of polyketides in heterologous hosts. *Microbiol Mol Biol Rev* **65**, 106-118.

Piel, J., Hertweck, C., Shipley, P. R., Hunt, D. M., Newman, M. S. & Moore, B. S. (2000a). Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster

from the marine isolate "*Streptomyces maritimus*": evidence for the derailment of an aromatic polyketide synthase. *Chem Biol* **7**, 943-955.

Piel, J., Hoang, K. & Moore, B. S. (2000b). Natural metabolic diversity encoded by the enterocin biosynthesis gene cluster. *J Am Chem Soc* **122**, 5415-5416.

Piepersberg, W. & Distler, J. (1997). Aminoglycosides and sugar components in other secondary metabolites. In *Biotechnology (2nd Edition)*, pp. 397-488. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Rawlings, B. J. (1999). Biosynthesis of polyketides (other than actinomycete macrolides). *Nat Prod Rep* **16**, 425-484.

Reysenbach, A. L., Wickham, G. S. & Pace, N. R. (1994). Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl Environ Microbiol* **60**, 2113-2119.

Ritacco, F. V., Graziani, E. I., Summers, M. Y., Zabriskie, T. M., Yu, K., Bernan, V. S., Carter, G. T. & Greenstein, M. (2005). Production of novel rapamycin analogs by precursor-directed biosynthesis. *Appl Environ Microbiol* **71**, 1971-1976.

Ritacco, F. V., Haltli, B., Janso, J. E., Greenstein, M. & Bernan, V. S. (2003). Dereplication of *Streptomyces* soil isolates and detection of specific biosynthetic genes using an automated ribotyping instrument. *J Ind Microbiol Biotechnol* **30**, 472-479.

Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., Tiong, C. L., Gilman, M., Osburne, M. S., Clardy, J., Handelsman, J. & Goodman, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* **66**, 2541-2547.

Saddler, G., O'Donnell, A. G. & Goodfellow, M. (1986). Fatty acid and menaquinone composition of *Streptomyces cyaneus* NCIB 9616 at different stages in the growth cycle. *Symposia Biologica Hungarica* **32**, 575-577.

Saddler, G. S., O'Donnell, A. G., Goodfellow, M. & Minnikin, D. E. (1987). SIMCA pattern recognition in the analysis of Streptomycete fatty acids. *J Gen Microbiol* **133**, 1137-1147.

Saddler, G. S., Tavecchia, P., Lociuro, S., Zanol, M., Colombo, L. & Selva, E. (1991). Analysis of madurose and other actinomycete whole cell sugars by gas chromatography. *J Microbiol Methods* **14**, 185-191.

Sanchez, S. & Demain, A. L. (2002). Metabolic regulation of fermentation processes. *Enzyme Microb Technol* **31**, 895-906.

Schatz, A., Bugie, E. & Waksman, S. A. (1944). Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc Soc Exp Biol Med* **55**, 66-69.

Schloss, P. D. & Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr Opin Biotechnol* **14**, 303-310.

Schwecke, T., Aparicio, J. F., Molnar, I., Konig, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J. & et al. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc Natl Acad Sci USA* **92**, 7839-7843.

Sehgal, S. N., Baker, H. & Vezina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. *J Antibiot (Tokyo)* **28**, 727-732.

Shen, B., Summers, R. G., Wendt-Pienkowski, E. & Hutchinson, C. R. (1995). The *Streptomyces glaucescens tcmKL* polyketide synthase and *tcmN* polyketide cyclase genes govern the size and shape of aromatic polyketides. *J Am Chem Soc* **117**, 6811-6821.

Shirling, E. B. & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313-340.

Sigmund, J. M., Clark, D. C., Rainey, F. A. & Anderson, A. S. (2003). Detection of eubacterial 3-hydroxy-3-methylglutaryl coenzyme A reductases from natural populations of actinomycetes. *Microb Ecol* **46**, 106-112.

Sosio, M., Bossi, E., Bianchi, A. & Donadio, S. (2000). Multiple peptide synthetase gene clusters in actinomycetes. *Mol Gen Genet* **264**, 213-221.

Stackebrandt, E., Liesack, W. & Witt, D. (1992). Ribosomal RNA and rDNA sequence analyses. *Gene* **115**, 255-260.

Stackebrandt, E., Rainey, F. A. & Ward-Rainey, N. L. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* **47**, 479-491.

Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R. & Liesack, W. (1991). Designation of streptomycete 16S and 23S rRNA-based target regions for oligonucleotide probes. *Appl Environ Microbiol* **57**, 1468-1477.

Strohl, W. R. (1997). Industrial antibiotics: today and the future. In *Drugs and the Pharmaceutical Sciences*, pp. 1-47. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Ueda, K., Tomaru, Y., Endoh, K. & Beppu, T. (1997). Stimulatory effect of copper on antibiotic production and morphological differentiation in *Streptomyces tanashiensis*. *J Antibiot* **50**, 693-695.

Van Wageningen, A. M. A., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., Jones, M., Jones, S. J. M. & Solenberg, P. J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem Biol* **5**, 155-162.

von Dohren, H. & Grafe, U. (1997). General aspects of secondary metabolism. In *Biotechnology (2nd Edition)*, pp. 1-55. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

von Dohren, H. & Kleinkauf, H. (1997). Enzymology of peptide synthetases. In *Drugs and the Pharmaceutical Sciences*, pp. 217-240. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Waksman, S. A. & Henrici, A. T. (1943). The nomenclature and classification of the actinomycetes. *J Bacteriol* **46**, 337-341.

Waksman, S. A. & Woodruff, H. B. (1940). Bacteriostatic and bactericidal substances produced by a soil *actinomyces*. *Proc Soc Exp Biol Med* **45**, 609-614.

Waksman, S. A. & Woodruff, H. B. (1942). Streptothricin, a new selective bacteriostatic and bactericidal agent, particularly active against gram-negative bacteria. *Proc Soc Exp Biol Med* **49**, 207-210.

Walsh, C. T. (2003). *Antibiotics: actions, origins, resistance*.ASM Press, Washington, D.C.

Walsh, C. T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805-1810.

Wawrik, B., Kerkhof, L., Zylstra, G. J. & Kukor, J. J. (2005). Identification of unique type II polyketide synthase genes in soil. *Appl Environ Microbiol* **71**: 2232-2238.

Wink, J. (2002). *The Actinomycetales: an order in the class of Actinobacteria important to the pharmaceutical industry- electronic manual*.CD-Rom Provided by Dr. Joachim Wink, Aventis Pharma Deutschland GmbH.

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

Xiang, L., Kalaitzis, J. A., Nilsen, G., Chen, L. & Moore, B. S. (2002). Mutational analysis of the enterocin favorskii biosynthetic rearrangement. *Org Lett* **4**, 957-960.

Xiang, L., Kalaitzis, J. A. & Moore, B. S. (2004). *EncM*, a versatile enterocin biosynthetic enzyme involved in Favorskii oxidative rearrangement, aldol condensation, and heterocycle-forming reactions. *Proc Natl Acad Sci USA* **101**, 15609-15614.

Xiang, L. & Moore, B. S. (2002). Inactivation, complementation, and heterologous expression of *encP*, a novel bacterial phenylalanine ammonia-lyase gene. *J Biol Chem* **277**, 32505-32509.

Xiang, L. & Moore, B. S. (2003). Characterization of benzoyl coenzyme A biosynthesis genes in the enterocin-producing bacterium "*Streptomyces maritimus*". *J Bacteriol* **185**, 399-404.

Yamada, Y., Nihira, T. & Sakuda, S. (1997). Butyrolactone autoregulators, inducers of virginiamycin in *Streptomyces virginiae*: their structures, biosynthesis, receptor proteins, and induction of virginiamycin biosynthesis. In *Drugs and the Pharmaceutical Sciences*, pp. 63-79. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Yu, T.-W., Shen, Y., McDaniel, R., Floss, H. G., Khosla, C., Hopwood, D. A. & Moore, B. S. (1998). Engineered biosynthesis of novel polyketides from *Streptomyces* spore pigment polyketide synthases. *J Am Chem Soc* **120**, 7749-7759.

Zabriskie, D. W., Armiger, W. B., Phillips, D. H. & Albano, P. A. (1980). *Traders' guide to fermentation media formulation*, Fourth Printing, 1994 edn.Traders Protein, Memphis, Tennessee.

Chapter 3

Molecular and phenotypic comparison of phaeochromycin-producing strains of *Streptomyces phaeochromogenes* and *Streptomyces ederensis*.

Frank V. Ritacco[1] and Douglas E. Eveleigh[2].

[(1)] Wyeth Research, Pearl River, NY, and [(2)] School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ.

Abstract

Streptomyces strain LL-P018 produces the phaeochromycins, novel anti-inflammatory polyketides. This organism was identified as a strain of *Streptomyces phaeochromogenes* by physiological and genetic taxonomic analysis. In order to gain greater taxonomic perspective, LL-P018 was compared to related strains from major culture collections by 16S rRNA gene sequence, ribotype, HPLC-MS metabolite profile, and *rpoB* sequence. Using BioNumerics software, genetic and chemical fingerprint data were integrated *via* multivariate cluster analysis into a single, robust comparison. Based upon this analysis, strain LL-P018 is very closely related to the type strains of both *S. phaeochromogenes* and *Streptomyces ederensis*, indicating that these two types may in fact represent a single species. This novel comparative multi-cluster analysis is most useful for illuminating strain variation among streptomycetes, and for clarifying relationships between closely related species.

Introduction

Streptomycetes are abundant soil bacteria capable of producing a vast array of complex and biologically active secondary metabolites. They have been studied extensively over the past 50 years and are the source of a large and diverse range of medically important and highly valuable pharmaceutical products including antibacterial, antifungal, antiparasitic, anticancer and immunosuppressant drugs (Baltz, 1997, 2006; Challis & Hopwood, 2003; Chater & Bibb, 1997; Clardy *et al.*, 2006; Hansen, 1997; Jack *et al.*, 1997; Mankelow & Neilan, 2000; Nicas & Cooper, 1997; Piepersberg & Distler, 1997; Schatz *et al.*, 1944; Sehgal *et al.*, 1975; Strohl, 1997; Waksman & Woodruff, 1942; Walsh, 2003, 2004; Wink, 2002).

Streptomyces strain LL-P018 was isolated from a soil sample taken from the bank of the Aa river in Westerenger, Germany. This strain produces the phaeochromycins (Figure 3.1), novel polyketide inhibitors of MAPKAP-2 (MK-2) (Graziani *et al.*, 2005). MK-2 is a kinase involved in the regulation of tumor necrosis factor (TNF-α) biosynthesis, and is a promising target in the search for new anti-inflammatory drugs (Kotlyarov *et al.*, 1999). The phaeochromycins are analogous to intermediates and shunt products of other type II polyketide biosynthetic pathways, differing primarily in the presence of a characteristic propyl side chain, possibly derived from a butyrate starter unit (Graziani *et al.*, 2005).

Strain LL-P018 was taxonomically identified using a combination of traditional and molecular biological techniques. While 16S rRNA gene sequencing and classical

taxonomic methods (such as morphological observation, carbohydrate utilization, sporulation and pigment production) are often sufficient for species identification, more sensitive methods are needed to detect strain variation within a species (Anderson & Wellington, 2001; Anzai *et al.*, 1997; Ritacco *et al.*, 2003). In order to determine the uniqueness of this strain with regard to its production of novel secondary metabolites, genetic and chemical fingerprinting methods were used to compare LL-P018 to fifteen closely related strains from major culture collections. Genetic fingerprinting of all strains was performed using the RiboPrinter®, an automated ribotyping instrument capable of discriminating between strains of *Streptomyces* (Pfaller *et al.*, 1996; Ritacco *et al.*, 2003). For comparison of chemical fingerprints, all strains were grown in three liquid media under identical conditions. Culture extracts were analyzed by HPLC-MS. BioNumerics® software (Applied Maths, Austin, TX) was used to integrate RiboPrint® patterns and multiple HPLC chromatograms in a single cluster analysis, allowing the simultaneous comparison of molecular typing data and fermentation metabolite profiles. This novel approach to classification is most useful in clarifying relationships between closely related strains and species.

Materials and Methods

Strains

Streptomyces strain LL-P018 has been previously documented (Graziani *et al.*, 2005) and was maintained at Wyeth Research, Pearl River, NY. Comparative strains were obtained from the Agricultural Research Service (ARS) Culture Collection, Peoria, IL and the American Type Culture Collection (ATCC), Manassas, VA (Table 3.1).

Because multiple "type" strains for *S. phaeochromogenes* exist in these and other culture collections, three of these "type" strains (NRRL B-1248^T, ATCC 3338^T and ATCC 23945^T) were used for comparison in this study. *S. phaeochromogenes* was first described as *Actinomyces phaeochromogenus* by H.J. Conn in 1917, but the oldest existing representative of this type is ATCC 3338^T, which was deposited by Selman Waksman in 1963 (ATCC technical services- personal communication). ATCC 3338^T is the official type strain for this species.

16S rRNA gene sequencing

Complete (>1400bp) 16S rRNA genes were amplified by polymerase chain reaction (PCR) using standard primers (Reysenbach *et al.*, 1994). PCR amplification was carried out in 100µL reaction volumes containing approximately 5–10 ng of genomic DNA, 1 mM each of the 8 FPL forward and 1492 RPL reverse primers and 50µL of Jumpstart PCR mix (Sigma, St. Louis, MO). The PCR was performed using a BioMetra T Gradient thermocycler (Whatman-BioMetra, Goettingen, Germany) as follows: one cycle of

denaturation at 95ºC for 2 min, followed by 35 cycles of 95ºC for 1 min, 55ºC for 45 s and 72ºC for 90 s, with one extension cycle at 72ºC for 5 min and a pause at 4ºC. The amplified PCR product was directly sequenced using an ABI 3700 sequencer with the ABI Prism DNA sequencing kit and Big Dye terminators version 3.0 (Applied Biosystems, Foster City, Calif.). Sequencing was performed on the PCR reaction using the 16S forward and reverse PCR primers described above, and the sequence was compared to the GenBank database by nucleotide BLAST search (Altschul *et al.*, 1997). Sequences were aligned using ClustalX (Thompson *et al.*, 1997). The phylogenetic tree was based on neighbor-joining analysis of 16S rRNA gene sequences with 1000 bootstrap replicates using TREECON software (Jukes & Cantor, 1969; Van de Peer & De Wachter, 1994). The 16S rRNA gene sequence of the evolutionarily distant species *Actinoplanes utahensis* IFO-13244^{T} was chosen as an outgroup and used to root the phylogenetic tree.

Carbohydrate utilization, pigmentation, and sporulation

Carbohydrate utilization, pigment production, and sporulation were observed according to the methods defined by the International *Streptomyces* Project (Shirling & Gottlieb, 1966). For analysis of carbohydrate utilization, agar media containing appropriate carbohydrate sources (4mL) was added to the wells of a 12-well Falcon culture plate (35-3043, Becton-Dickinson, Franklin Lakes, NJ) and inoculated with 25µL of cells prepared and washed as described (Shirling & Gottlieb, 1966). For observation of sporulation and pigment production, 50µL washed cells were inoculated to ISP sporulation agars and

grown under defined conditions (Shirling & Gottlieb, 1966). Growth, sporulation, and pigment production were observed after 7, 14, and 21 days.

Ribotyping

Ribotype analysis was performed using the RiboPrinter Microbial Characterization System (DuPont Qualicon, Wilmington, DE) as previously described (Ritacco *et al.*, 2003). Two 10µL loopfuls containing 2–5 colonies were scraped from polycarbonate membranes that were overlaid upon agar plates. Colonies were ground with a sterile, motorized pestle in 200µL of RiboPrinter sample buffer (DuPont Qualicon, Wilmington, DE). Forty microliters of each ground cell suspension was heat-treated in the RiboPrinter heat-treatment station for approximately 20 min. Lysis buffers (DuPont Qualicon, Wilmington, DE) were then added to each sample and the samples were loaded into the RiboPrinter. Disposable reagents, including *Pvu*II, DNA preparative enzymes and reagents, probe, conjugate and substrate solutions, agarose gel and membrane were loaded into the RiboPrinter, according to the manufacturer's instructions (Qualicon) and the methods of Bruce, *et al.* (Bruce, 1996). Genomic DNA obtained from cell lysates was digested with *Pvu*II, and chromosomal DNA fragments were hybridized with the ribosomal probe provided by the manufacturer (DuPont Qualicon, Wilmington, DE).

Culture conditions

All cultures were grown in three liquid media. Liquid medium 1 contained (grams per liter): $MgSO_4 \cdot 7H_2O$ (0.5), KCl (0.5), K_2HPO_4 (2.5), NaCl (5), agar (0.4), glycerol (10), soy peptone (5). Liquid medium 2 contained (grams per liter): D-mannitol (10), Bacto-

peptone (Difco/Becton-Dickinson, Franklin Lakes, NJ) (1), soluble starch (Difco/Becton-Dickinson, Franklin Lakes, NJ) (15), $MgSO_4 \cdot 7H_2O$ (0.1), KH_2PO_4 (0.5), and $CaCO_3$ (1). Liquid medium 3 contained (grams per liter): D-glucose (1), D-mannitol (10), Bacto-peptone (Difco/Becton-Dickinson, Franklin Lakes, NJ) (1), soluble starch (Difco/Becton-Dickinson, Franklin Lakes, NJ) (15), $MgSO_4 \cdot 7H_2O$ (0.1), KH_2PO_4 (0.5), and $CaCO_3$ (1). Cultures were grown at 28ºC and 200 rpm for 6 days. One milliliter (1mL) of whole broth was extracted in 1mL ethyl acetate (1:1) and centrifuged to separate phases. The organic phase was dried to completion and the product was resolubilized in 100 µL methanol, resulting in a 10-fold final concentration. Reverse phase HPLC of extracts was performed using an Agilent model 1100 liquid chromatograph with photodiode array detection, coupled to an Agilent LC-MSD ion trap mass spectrometer fitted with an electrospray ionization source. Extracts were resolved by reverse phase chromatography using a Zorbax Eclipse XDB-C18 column (3.0 x 75mm, 3.5µm particle size) with a mobile phase of 0.025% formic acid in water (solvent A) and 0.025% formic acid in acetonitrile (solvent B). For elution, a linear gradient from 5% A to 95% B in 15 min was used, holding at 95% acetonitrile for 10 min, with a flow rate of 0.8 mL/min. A split was used to divert 25% of the flow into the mass spectrometer. UV detection was performed using a photodiode array with a scan range of 200-600 nm. The MS electrospray was performed in positive mode with a scan range of 200-2,000 m/z.

Cluster analysis of ribotype data and HPLC chromatograms

Cluster analysis of RiboPrints and HPLC-MS chromatograms was performed using BioNumerics software (Applied Maths, Austin, Tex). RiboPrint patterns were exported

from the RiboPrinter database as text files and uploaded into a BioNumerics database using a custom import script (BNScripts60) designed by software engineers at Applied Maths (Austin, TX). HPLC chromatograms were exported from Agilent ChemStation software as comma-delimited Excel files (Microsoft) and uploaded into the BioNumerics database using another custom import script (import_curves_background). Using this script, each exported chromatogram file is uploaded into BioNumerics along with a "blank" chromatogram to be used for background subtraction. During import into BioNumerics, an HPLC chromatogram derived from an uninoculated media blank was subtracted from each sample chromatogram to eliminate unwanted and non-specific medium components from the resulting comparisons. RiboPrint Patterns and HPLC chromatograms were analyzed as densitometric curves and compared using the Pearson product-moment correlation (Pearson, 1926; Vauterin & Vauterin, 2006). Hierarchical cluster analysis of Riboprints and HPLC chromatograms was performed using the unweighted pair group method using arithmetic averages (UPGMA). For composite data sets comprised of data from multiple experiments, RiboPrint patterns and metabolite profiles were equally weighted, and the resemblance matrices of the individual data sets were averaged (Vauterin & Vauterin, 2006).

***rpoB* sequencing**

354bp Ribosomal polymerase β-subunit (*rpoB*) genes were amplified by polymerase chain reaction (PCR) using defined primers (Kim *et al.*, 2004). PCR amplification was carried out in 100μL reaction volumes containing approximately 5–10 ng of genomic DNA, 1 mM each of the SRPOF1 forward and SRPOR1 reverse primers (Kim *et al.*,

2004) and 50µL of Jumpstart RedTaq PCR mix (Sigma, St. Louis, MO). The PCR was performed under conditions recommended by Dr. David Labeda (USDA, Peoria, IL, personal communication) using a BioMetra T Gradient thermocycler (Whatman-BioMetra, Goettingen, Germany) as follows: one cycle of denaturation at 95ºC for 5 min, followed by 40 cycles of 95ºC for 30 seconds, 62ºC for 30 seconds and 72ºC for 45 seconds, with one extension cycle at 72ºC for 5 min and a pause at 4ºC. The amplified PCR product was directly sequenced using an ABI 3700 sequencing machine with the ABI Prism DNA sequencing kit and Big Dye terminators version 3.0 (Applied Biosystems, Foster City, Calif.). Sequencing was performed on the PCR reaction using the forward and reverse PCR primers described above, and the sequence was compared to the GenBank database by nucleotide BLAST search (Altschul *et al.*, 1997). Sequences were aligned using ClustalX (Thompson *et al.*, 1997). The phylogenetic tree was based on neighbor-joining analysis of *rpoB* sequences with 1000 bootstrap replicates using TREECON software (Jukes & Cantor, 1969; Van de Peer & De Wachter, 1994).

Results

16S rRNA gene sequencing

A phylogenetic tree depicting comparison of all strains by 16S rRNA gene sequence showed four clusters that were at least 99% similar by 16S rRNA gene sequence (Figure 3.2). The largest of these, Cluster 1, contained seven strains including *Streptomyces* strain LL-P018. The 16S rRNA gene of strain LL-P018 was 99.5% similar to that of the type strains of *S. phaeochromogenes* included in this study (Figure 3.2). Two other

strains of *S. phaeochromogenes* (NRRL B-5478 and ISP-5170) and *S. ederensis* NRRL B-8146^T were also clustered with this highly similar group. Of the strains in Cluster 1, five (*S. phaeochromogenes* ATCC 23945^T, S. *phaeochromogenes* NRRL B-5478, strain LL-P018, *S. phaeochromogenes* ISP-5170, and *S. ederensis* NRRL B-8146^T) exhibited heterogeneity in the hypervariable α-region of the 16S rRNA gene (Figure 3.3), indicating slight sequence differences between multiple copies of the gene within each organism (Ueda *et al.*, 1999).

Streptomyces tauricus NRRL B-12497^T was a divergent offshoot from Cluster 1, sharing 98.3% sequence similarity with the *S. phaeochromogenes* type strains (Figure 3.2). The remaining 8 strains were highly divergent from the *S. phaeochromogenes* cluster. Comparison of these sequences to the GenBank database suggested that these strains are more closely related to other species of *Streptomyces* than *S. phaeochromogenes* (Figure 3.2).

Cluster 2 contained NRRL B-5333 and NRRL B-3559, whose 16S rRNA gene sequences were 99.8% similar and most closely matched that of a *S. viridobrunneus* type strain in GenBank.

Strains NRRL WC-3776 and NRRL B-2031 comprised Cluster 3. These strains are 99.9% similar by 16S rRNA gene sequence, and are most closely related to *S. viridobrunneus*.

The fourth cluster contained strains NRRL B-2123 and NRRL B-1517, which shared 99.1% sequence identity and most closely resembled the *S. melanogenes* type strain.

Three strains, *S. tauricus* NRRL B-12497^T, *S. phaeochromogenes* NRRL B-16392, and *S. phaeochromogenes* NRRL B-1131 did not share greater than 99% sequence identity with any other strain and were not assigned to a cluster.

Carbohydrate utilization

Carbohydrate utilization patterns confirmed identification of strain LL-P018 as a strain of *S. phaeochromogenes* (Table 3.2). All carbon sources tested (except cellulose) were utilized readily by strain LL-P018, *S. phaeochromogenes* ATCC 3338^T, *S. ederensis* NRRL B-8146^T, and *S. tauricus* NRRL B-12497^T. Slight growth on cellulose (microcrystalline, Sigma, St. Louis, MO) was observed only for strain LL-P018 and *S. ederensis* NRRL B-8146^T.

Sporulation and pigment production

Observations of sporulation and pigment production on ISP agars are listed in Table 3.3. Strain LL-P018 was highly similar to ATCC 3338^T, supporting placement of strain LL-P018 within the species *S. phaeochromogenes*. A strong similarity between the *S. ederensis* type strain and the *S. phaeochromogenes* type strain was also observed.

Ribotyping

Three "type" strains of *S. phaeochromogenes* (NRRL B-1248^T, ATCC 3338^T and ATCC 23945^T) were greater than 95% similar to one another by ribotype, confirming that these are identical strains (Figure 3.4). Strain LL-P018 grouped closely with these strains, forming a distinct branch 92% similar to the type strain cluster. This variation is clearly due to the presence of an approximately 11kb band in the ribotype fingerprint of strain LL-P018 which is not seen in those of the type strains, indicating genomic divergence and therefore strain variation. The *S. ederensis* type strain, NRRL B-8146^T, was 97.4% similar to *S. phaeochromogenes* strain NRRL B-5478 by ribotype. These strains both clustered with strain LL-P018 and the *S. phaeochromogenes* type strains by 16S rRNA gene sequence comparison (Figure 3.2, "Cluster 1"), but were only 80% similar to these strains by ribotype and therefore formed a distinct ribotype cluster (Figure 3.4). *S. phaeochromogenes* ISP-5170, which also grouped in 16S rRNA gene cluster 1 (Figure 3.2), was very divergent from the group by ribotype, sharing only 21%-34% similarity to other strains in this group (Figure 3.4).

The remaining 8 strains, which were not closely related to *S. phaeochromogenes* or *S. ederensis* by 16S rRNA gene sequence or ribotype (3.s 2 and 4), exhibited ribotype strain variation within their respective 16S rRNA gene clusters. Strains NRRL B-5333 and NRRL B-3559 (Figure 3.2, "Cluster 2"), which were 99.8% identical in 16S rRNA gene sequence, shared only 75.9% similarity by ribotype (Figure 3.4). Strains NRRL WC-3776 and NRRL B-2031 (Figure 3.2, "Cluster 3") were 99.9% similar in 16S rRNA gene sequence but only 13.9% similar by ribotype (Figure 3.4). Strains NRRL B-2123 and

NRRL B-1517 (Figure 3.2, "Cluster 4") were only 79.4% similar by ribotype (Figure 3.4), despite 99.1% similarity in 16S rRNA gene sequence (Figure 3.2). Based on 16S rRNA similarity values, these strain pairs clearly represent species-level clusters. However, the divergence in ribotype patterns just described suggests subspecies or strain-level variation within these clusters. These apparent strain variations were investigated further by comparison of metabolite profiles, and will be discussed in detail in this chapter.

Metabolite profiling

Strain LL-P018 and the *S. phaeochromogenes* type strains were all highly similar by metabolite profile (Figure 3.5). In liquid media 1 and 2, these strains produced only the phaeochromycins (Figure 3.6). In liquid medium 3, however, these strains produced only the related compound alnumycin. Metabolite profiles for these strains correlated well with ribotype comparisons and confirmed a high degree of similarity between Strain LL-P018 and *S. phaeochromogenes.* For example, in a composite analysis combining metabolite profiles from all three media, these four strains formed a cluster sharing 83-97% similarity to one another (Figure 3.5).

The *S. ederensis* type strain, NRRL B-8146^T was very similar to *S. phaeochromogenes* NRRL B-5478 (88.3%) and *S. phaeochromogenes* ISP-5170 (92.2%) by composite metabolite profile (Figure 3.5). These three strains grouped closely with strain LL-P018 and the *S. phaeochromogenes* type strains in 16S rRNA gene sequence (Figure 3.2, "Cluster 1") but showed distinct differences in metabolite profile (Figure 3.5). All of the

strains in 16S rRNA gene cluster 1 (Figure 3.2) produced the phaeochromycins in liquid medium 1 (Figure 3.6) and alnumycin in liquid medium 3. However, differences in metabolite production among members of this group were observed in liquid medium 2. In this medium, NRRL B-8146^T, NRRL B-5478 and ISP-5170 produced alnumycin. In contrast, strain LL-P018 and the *S. phaeochromogenes* type strains produced the phaeochromycins (Figure 3.5). As a result, strains NRRL B-8146^T, NRRL B-5478 and ISP-5170 formed a separate cluster in the composite HPLC comparison, only 64.6% similar in metabolite profile to the cluster containing strain LL-P018 and the *S. phaeochromogenes* type strains (Figure 3.5).

Strains NRRL B-5333 and NRRL B-3559 (Figure 3.2, "Cluster 2") were only 52% similar by composite metabolite profile (Figure 3.5). This result supports the strain variation suggested by ribotype comparison (76% similarity) (Figure 3.4). Very low metabolite similarity (6.9%) between strains NRRL WC-3776 and NRRL B-2031 (Figure 3.2, "Cluster 3", and Figure 3.5) was in agreement with the low ribotype similarity between these strains (13.9%), but stood in stark comparison to the 99% similarity in their 16S rRNA gene sequences (Figure 3.2). Strains NRRL B-2123 and NRRL B-1517 (Figure 3.2, "Cluster 4") were 69.6% similar by metabolite profile (Figure 3.5), sharing several major metabolite peaks but showing differences in the presence and intensity of many analytes as well.

Combined ribotype and metabolite profile comparison

Integration of genetic and chemical fingerprint data into a single comparison yielded a robust analysis combining elements of each method (Figure 3.7). For this analysis, one ribotype pattern and three HPLC metabolite profiles were combined with equal weighting. The seven strains comprising 16S rRNA gene cluster 1 (Figure 3.2) formed two distinct clusters in this composite analysis (Figure 3.7). One contained strain LL-P018 and the three *S. phaeochromogenes* type strains, which were all highly similar by ribotype and metabolite profile (Figures 3.4, 3.5 and 3.7). The other contained *S. ederensis* NRRL B-8146^{T}, *S. phaeochromogenes* NRRL B-5478 and *S. phaeochromogenes* ISP-5170 (Figures 3.2 and 3.7). This group differed from strain LL-P018 and the *S. phaeochromogenes* type strains in both ribotype (Figure 3.4) and in production of the phaeochromycins in liquid medium 2 (Figure 3.5). Strain ISP-5170, which was highly divergent from the other Cluster 1 strains by ribotype (Figure 3.4), rejoined the group in this composite analysis due to overall similarities in metabolite profile (Figure 3.7). Due to its unique ribotype pattern, however, ISP-5170 still formed a distinct branch within the cluster.

Strains NRRL B-5333 and NRRL B-3559 (Figure 3.2, "Cluster 2") were only 58% similar in this combined analysis due to significant variations in both ribotype and metabolite profile (Figure 3.7). Strains NRRL WC-3776 and NRRL B-2031 (Figure 3.2, "Cluster 3") were only 8.7% similar and did not cluster at all in this analysis (Figure 3.7), despite a high degree of 16S rRNA gene similarity (Figure 3.2). Strains NRRL B-2123 and NRRL B-1517 (Figure 3.2, "Cluster 4") were 72.6% similar in the combined analysis

(Figure 3.7), highlighting a degree of strain variation between these closely related organisms (see Discussion).

RpoB sequencing

Strain LL-P018 was compared to all strains from 16S rRNA gene cluster 1 (Figure 3.2) as well as *S. tauricus* NRRL B-12497^T by *rpoB* sequence (354bp). A phylogenetic tree depicting this comparison is shown in Figure 3.8. All of the strains from 16S rRNA gene cluster 1 (Figure 3.2) were 99.7-100% similar in *rpoB* sequence (Figure 3.8), consistent with intraspecies similarity values reported for *Streptomyces* species (Kim *et al.*, 2004). The *S. ederensis* type strain, NRRL B-8146^T, was 100% identical to *S. phaeochromogenes* ATCC 3338^T, *S. phaeochromogenes* ATCC 23945^T, and *S. phaeochromogenes* NRRL B-1248^T (Figure 3.8), further suggesting that *S. ederensis* and *S. phaeochromogenes* represent a single species. *Streptomyces tauricus* NRRL B-12497T shared 95.2% *rpoB* sequence similarity with the strains from 16S rRNA gene cluster 1 (Figures 3.2 and 3.8), consistent with previously reported interspecies similarity values for species within the genus *Streptomyces* (Kim *et al.*, 2004).

Discussion

Type strains of *S. phaeochromogenes* and the closely related species *S. ederensis* and *S. tauricus* were obtained for comparison to strain LL-P018. Comparison of these strains by 16S rRNA gene sequence, morphological analysis, sporulation, pigmentation, and carbohydrate utilization supported placement of LL-P018 within the species *S.*

phaeochromogenes, and also suggested a high degree of similarity between the type strains for *S. phaeochromogenes* and *S. ederensis* (Figure 3.2, Tables 3.2 and 3.3). In order to assess the uniqueness of strain LL-P018 among members of this species, 10 strains identified as *S. phaeochromogenes* were obtained from the USDA-ARS (NRRL) culture collection (Table 3.1). These strains, along with type strains of *S. phaeochromogenes, S. ederensis* and *S. tauricus* were compared to strain LL-P018 by 16S rRNA gene analysis as well as genetic and chemical fingerprinting methods.

All strains were first compared by 16S rRNA gene sequence analysis (Figure 3.2). Strain LL-P018 was 99.5% similar to the type strains of *S. phaeochromogenes* (Figure 3.2). Two other strains of *S. phaeochromogenes*, NRRL B-5478 and ISP-5170 were also highly similar to the *S. phaeochromogenes* type strains (99.5% and 100%, respectively, Figure 3.2). Interestingly, the type strain for *S. ederensis* also falls within this cluster and shares 99.6% 16S rRNA gene sequence similarity to the *S. phaeochromogenes* type strains (Figure 3.2). This result was consistent with comparison of *rpoB* sequences (Figure 3.8), as well as morphology, sporulation and carbohydrate utilization profiles of these strains, which were all highly similar (Tables 3.2 and 3.3). These results, along with ribotyping and metabolite profiling data (discussion below) indicate that *S. phaeochromogenes* and *S. ederensis* are highly similar, and in fact represent a single species. We therefore recommend reclassifying *S. ederensis* as a strain of *S. phaeochromogenes*, distinct from the type by ribotype and variation in metabolite profile.

Of the 16 strains compared by 16S rRNA gene sequence, 7 (including strain LL-P018) formed a very tight cluster with 99%-100% sequence identity (Figure 3.2, "Cluster 1"). Thorough analysis of sequencing chromatograms revealed that slight 16S rRNA gene sequence variation between these highly related strains is due in part to heterogeneity in the hypervariable region of the gene (Figure 3.3). Sequence heterogeneity between multiple copies of the 16S rRNA gene within a single strain of *Streptomyces* has been reported, and may occur in up to 6.9% of species within this genus (Ueda *et al.*, 1999). This heterogeneity occurs within the highly variable α-region of the gene, among the nucleotides corresponding to bases 179-197 of the *S. ambofaciens* genome, which form a stem-loop in the secondary structure of the RNA (Ueda *et al.*, 1999).

Sequence heterogeneity can be observed if the 16S rRNA gene is directly sequenced from a PCR product, in which case multiple copies of the gene are amplified using the primer set and represented in the sequencing reaction. Analysis of sequencing chromatograms can reveal the presence of multiple nucleotide signals at a single base location. This heterogeneity can be taken into consideration when comparing closely related strains, especially when sequence variation between strains occurs in the stem-loop region. Unfortunately, this heterogeneity is not always represented in sequence data deposited in GenBank. This may be due to the use of a sequencing method where the PCR product is cloned prior to sequencing. By such a method only one copy of the gene is sequenced, leading to a much cleaner sequencing reaction but missing evidence of sequence heterogeneity. Another possibility is that investigators may attempt to "clean up" unclear sequence data, by selecting the stronger of two concurrent nucleotide signals

as the correct base identification. This would also ignore the presence of multiple 16S rRNA gene sequences within the organism, and therefore yield an incomplete data set for comparison.

Eight strains deposited in the NRRL culture collection by various investigators as *S. phaeochromogenes* were found to be more closely related to other *Streptomyces* type strains by 16S rRNA gene sequence comparison (Figure 3.2). These strains were originally chosen for comparison to strain LL-P018 based on their species identification in the NRRL database. It is likely, however, that these strains were not fully characterized taxonomically prior to deposition. It should be mentioned that only type strains can be assumed thoroughly characterized and identified. Any other strains which may be available from a culture collection, or which may arise as results of a GenBank query, do not necessarily have a valid and reliable taxonomic identification. However, it should also be noted that though eight strains are now realized not to be closely related to *S. phaeochromogenes* or strain LL-P018 (Figure 3.2), they still proved valuable in the development of the genetic and chemical fingerprinting tools that were used in this study.

Evaluation of intraspecific relationships and strain variation below the species level requires methods more sensitive than 16S rRNA gene sequence analysis. Various genetic fingerprinting methods exist that have been used to discriminate between closely related *Streptomyces* strains, or to identify strain diversity in a group of environmental isolates (Anderson & Wellington, 2001; Anzai *et al.*, 1997; Lanoot *et al.*, 2005; Ritacco *et al.*, 2003). These molecular typing methods provide a "snapshot" of the genome, and can be

highly sensitive to slight differences in DNA sequence. As a result, two strains with very similar DNA fingerprints are typically considered to be identical isolates. However, slight variations in DNA fingerprint between two strains do not necessarily reflect differences in secondary metabolite profile or other readily observable phenotypes (Ritacco *et al.*, 2003). In order to take into account both genetic and metabolic strain variation, chemical fingerprinting of culture broth extracts (metabolite profiling) was used along with ribotyping, a well-established genetic fingerprinting method (Bruce, 1996; Ritacco *et al.*, 2003), to assess the similarity between strain LL-P018 and the related organisms selected for comparison.

Results of ribotype analysis suggested a close relationship between strain LL-P018 and the type strains of *S. phaeochromogenes* and *S. ederensis* (Figure 3.4), which had all grouped together in 16S rRNA gene cluster 1 (Figure 3.2). Three type strains of *S. phaeochromogenes* were greater than 96% similar to one another by ribotype (Figure 3.4), confirming that these are identical strains. Strain LL-P018 also grouped closely with these strains, forming a distinct branch 92% similar to the *S. phaeochromogenes* type strain cluster (Figure 3.4). This variation is clearly due to the presence of an approximately 11kb band in the ribotype fingerprint of strain LL-P018 which is not seen in those of the type strains. This degree of ribotype variation indicates the presence of some genomic divergence, but does not necessarily suggest obvious phenotypic distinctions. RiboPrint® patterns generated by the RiboPrinter® are typically considered identical if they are at least 93% similar, although a 92% similarity cutoff has been described for use with *Streptomyces* strains, which generate less RiboPrint® bands than

typical clinical isolates (Pfaller *et al.*, 1996; Ritacco *et al.*, 2003). A high degree of similarity between these four strains was further confirmed by metabolite profiling (Figures 3.5 and 3.6), as these strains produce the same suite of compounds under similar culture conditions. In a composite cluster analysis of ribotype and metabolite profiles, these four strains were greater than 89% similar to each other (Figure 3.7).

Two other strains from 16S rRNA gene cluster 1 (Figure 3.2), *S. ederensis* NRRL B-8146^T and *S. phaeochromogenes* strain NRRL B-5478, were 97% similar to one another by ribotype (Figure 3.4). These two strains were greater than 99% similar in 16S rRNA gene sequence to strain LL-P018 and the *S. phaeochromogenes* type strains (Figure 3.2, "Cluster 1"), but were only 80% similar to these strains by ribotype and therefore formed a distinct cluster in the ribotype analysis (Figure 3.4). Metabolite profiling confirmed this strain variation (Figure 3.5). Overall, these two strains produce the same family of compounds as strain LL-P018 and the *S. phaeochromogenes* type strains, but the profile of metabolites varies between the two strain clusters under certain culture conditions. This may indicate a difference in regulation of biosynthesis between these strain groups. Strain ISP-5170, also present in 16S rRNA gene cluster 1 (Figure 3.2), was even more divergent from the group by ribotype comparison, sharing less than 35% similarity with the other strains in the cluster (Figure 3.4). While this degree of ribotype variation would appear to indicate significant strain variation, metabolite profiling showed that this strain behaves similarly to NRRL B-8146^T and NRRL B-5478 in terms of phaeochromycin and alnumycin production (Figures 3.5 and 3.6). In the composite analysis combining ribotype and metabolite profiles, Strain ISP-5170 clustered most closely with NRRL B-

8146^T and NRRL B-5478, but formed a distinct branch within this group due to its unique ribotype pattern (Figure 3.7). This confirms the prediction that although ribotyping is a good indicator of strain variation, changes in DNA sequence that result in ribotype variation between two strains may not always affect obvious phenotypic traits such as secondary metabolite profiles (Ritacco *et al.*, 2003). This also confirms the necessity of combining genetic and phenotypic analyses in intraspecific and strain comparisons.

Of the strains that were not closely related to *S. phaeochromogenes* or *S. ederensis*, several showed strain variation within a species-level cluster. 16S rRNA gene clusters 2 and 4 (Figure 3.2) both exhibited variations in metabolite profile that correlated well with ribotype variation (Figure 3.7). In the case of 16S rRNA gene cluster 3 (Figure 3.2), strains NRRL WC-3776 and NRRL B-2031, although highly similar in 16S sequence, did not cluster at all in the combined analysis and appear to represent very different strains (Figure 3.7).

These results confirm the ability of genetic and chemical fingerprinting methods to illustrate strain variation among closely related strains, and highlight the value of this polyphasic approach to strain delineation.

There are several issues that must be considered with regards to metabolite profiling as described in this study. A critical factor in comparing HPLC chromatograms from culture extracts is assuring environmental and chromatographic consistency among

samples. In this study, strains were inoculated and cultured simultaneously under identical conditions to minimize variation due to environmental factors. Chromatographic equipment was regularly calibrated and standards were run intermittently among samples to assure that no retention time "drift" was occurring. These types of issues can be carefully controlled and monitored for the sake of a study such as this one, where a small set of samples is being compared. However, these issues may confound attempts to build a database of chemical fingerprints acquired over time, with new samples being added to the set periodically. In this case, careful normalization with the use of internal or external standards would be necessary to control changes in chromatography over time. Standardization of culture conditions and environmental factors would also be critical for the generation of reliable and reproducible data. It is for precisely these reasons that genetic fingerprinting is a very desirable method for strain delineation, as the genome of an organism is much less sensitive to this type of variability than metabolite production.

Also of importance in metabolite profiling is the processing of data. In this study, BioNumerics software was used for comparison of metabolite profiles. Typical metabolite profiles derived from HPLC analysis of culture extracts contain some artifacts that may not be specific to the strain being analyzed. Strong solvent peaks occur early in the chromatographic profile, while non-polar metabolites and media components may cause high background signal late in the chromatogram. Thus, in this study a 15-minute chromatogram was generated, but data generated in the first 2 minutes and final 3 minutes of the chromatogram were not included in the comparison. This eliminated

nonspecific peaks and areas of high background from the comparison. Also, background subtraction was performed on each chromatogram in the analysis. This was done by subtraction of a chromatogram of an uninoculated media blank from each sample chromatogram. The main effects of background subtraction on a metabolite profile are the elimination of peaks derived from medium components, overall reduction of background signal, and relative intensification of new peaks generated by the fermentation. While this may appear to have obvious benefit in profile analysis, issues such as the generation of "negative peaks" complicate this issue.

Negative peaks may be generated by the subtraction of media components that have already been utilized by a strain. Because the organism has consumed the medium component, the corresponding peak is not present in the sample chromatogram. Therefore, subtraction of the "blank" medium extract containing the peak results in negative values in the culture extract chromatogram. For this study, negative peaks were eliminated using BioNumerics. Although such negative peaks may be considered an important part of an organism's metabolite profile, in this study it was assumed that the consumption of a specific medium component is a characteristic that may be shared by multiple organisms, regardless of their relatedness at any taxonomic level. Negative peaks were therefore considered to be less strain-specific than peaks representing new metabolites produced, especially since the issue of secondary metabolite production is of relevance to this study. All of these considerations regarding the many possible approaches to metabolite profiling are subject to the opinions and the research objectives

of the investigator, and certainly a more thorough evaluation of these methods with a larger number of strains may shed more light on the best approaches to this task.

Streptomyces phaeochromogenes was first described by H.J. Conn in 1917 as *Actinomyces phaeochromogenus* (Conn, 1917). Conn states:

> "Description of species of Actinomycetes is at present very difficult. The literature abounds with descriptions that have become invalidated as better methods of study have been developed."

Conn's statement is still relevant and meaningful today, 90 years after it was written. Understanding the genetic and phenotypic relationships between these soil microorganisms is as complex as it was in Conn's day, even using current technologies. Clearly, no single taxonomic method has yet been described which can define all of the inter- and intra-specific relationships between members of the species *Streptomyces* or the Actinomycetes as a whole. Polyphasic methods incorporating genotypic and phenotypic comparisons have long been necessary for strain and species group comparisons (Anderson & Wellington, 2001), and as new technologies improve our ability to visualize these relationships, we will better understand these organisms and the abundant bioactive compounds they produce. Of particular promise is the fact that the phaeochromycins, a family of novel anti-inflammatory polyketides produced by one of the earliest described members of the species *Streptomyces*, were only just discovered in 2005. This is evidence that these ubiquitous and highly productive soil organisms, which

have been studied at great length for more than a century (Beijerinck, 1900; Rullmann, 1895; Waksman, 1959; Waksman & Curtis, 1916), still have some surprises for us. Furthermore, with the implementation of new methods and technologies providing a "fresh" look at "old" strains, more and more of the hidden potential of *Streptomyces* will continue to be discovered.

Acknowledgements

The authors would like to acknowledge Kyle Kingsley (Applied Maths, Austin, TX) and David Labeda (United States Department of Agriculture, Peoria, IL) for assistance with taxonomic experiments, Jan Kieleczawa for DNA sequencing, and Leonard McDonald for chemistry advice and assistance.

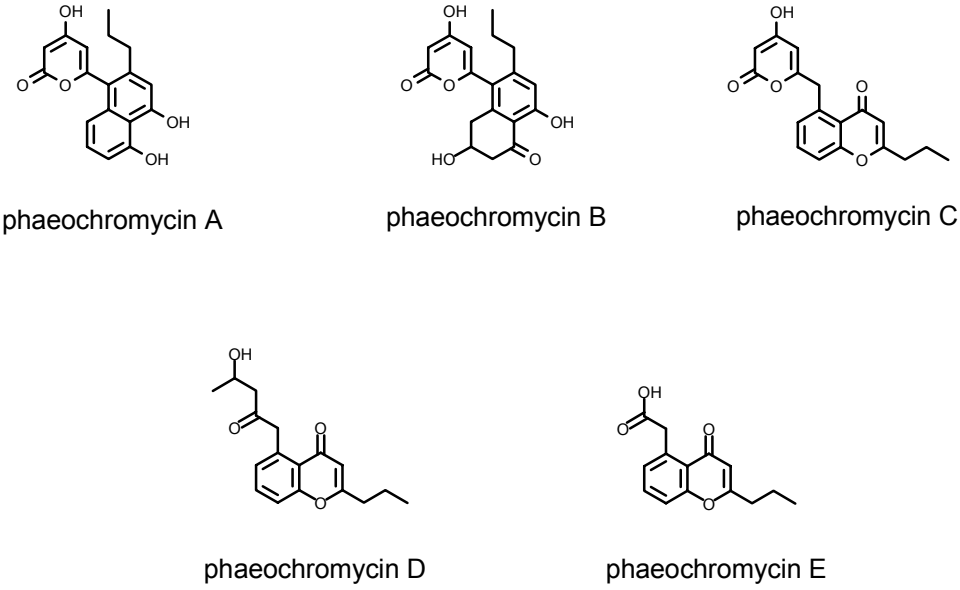


Figure 3.1. Polyketide structures of the phaeochromycins (Graziani *et al.*, 2005).

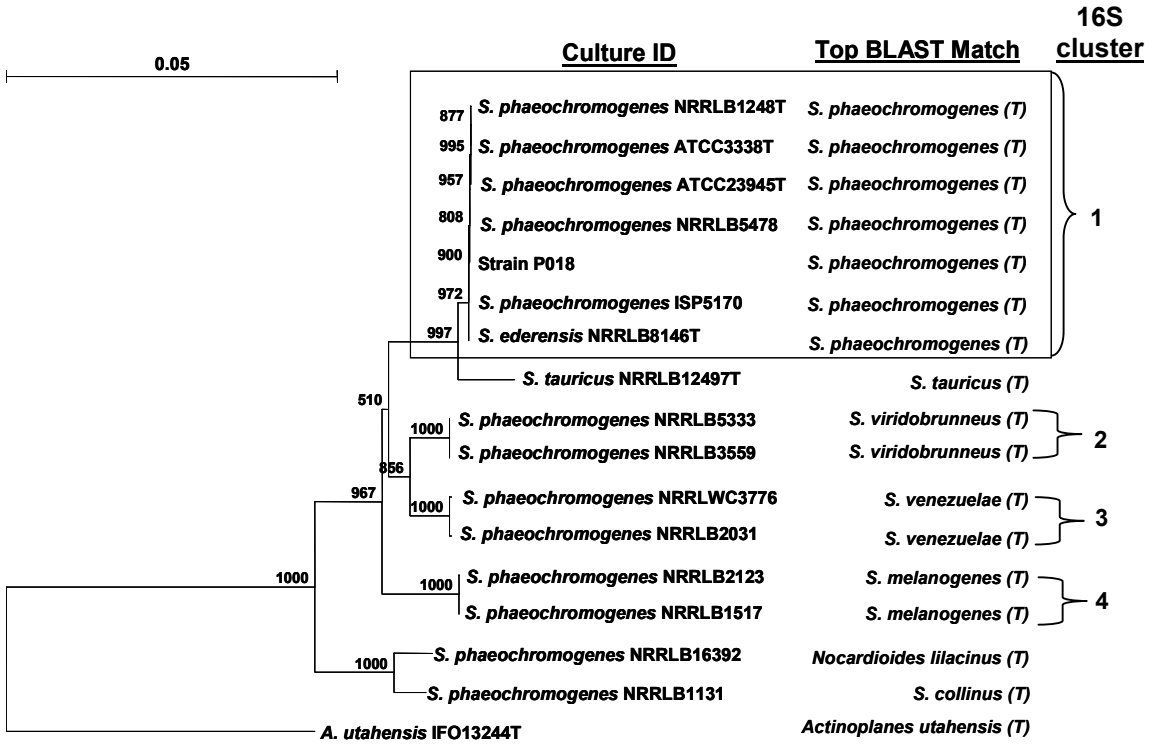


Figure 3.2. 16S rRNA sequence comparison of *S. phaeochromogenes* and closely related strains.

Complete (>1400bp) 16S ribosomal RNA (16S rRNA) gene sequences were aligned using ClustalX (Thompson *et al.*, 1997). The tree was constructed with TREECON software using the neighbor-joining method with 1000 bootstrap replicates (Jukes & Cantor, 1969; Van de Peer & De Wachter, 1994). Strains sharing greater than 99% sequence similarity were assigned a 16S rRNA gene cluster number for reference purposes. (T) = type strain.

	165				185
NRRL B-5478	CGGGGTCTAA	TACCGGATAA	CACTYSYMSA	GGCATCTSKG	KGGGTTGAA
Strain LL-P018	CGGGGTCTAA	TACCGGATAA	YACTYKYMSA	GGCATCTSKG	KGGGTTGAA
ISP-5170	CGGGGTCTAA	TACCGGATAA	CACTCSYSSM	SGCATSKSKG	KGGGTTGAA
NRRLB-8146^{T}	CGGGGTCTAA	TACCGGATAA	CACTCSYSSM	SGCATSKSKG	KGGGTTGAA
ATCC3338^{T}	CGGGGTCTAA	TACCGGATAA	CACTCGCCGA	GGCATCTCGG	TGGGTTGAA
NRRLB1248^{T}	CGGGGTCTAA	TACCGGATAA	CACTCGCCGA	GGCATCTCGG	TGGGTTGAA
ATCC23945^{T}	CGGGGTCTAA	TACCGGATAA	CACTCGCCSA	GGCATCTCGG	KGGGTTGAA

Figure 3.3. Sequence heterogeneity within the 16S rRNA genes of *S. phaeochromogenes* and closely related strains.

A portion of the 16S rRNA sequences of all strains from 16S rRNA gene cluster 1 (Figure 3.2) containing heterogeneous bases is shown. Sequences were aligned using ClustalX (Thompson *et al.*, 1997). Numbers shown refer to the corresponding nucleotide position in the LL-P018 16S rRNA gene sequence. Sequence heterogeneity codes: Y= C or T; K= G or T; S= C or G; M= A or C.

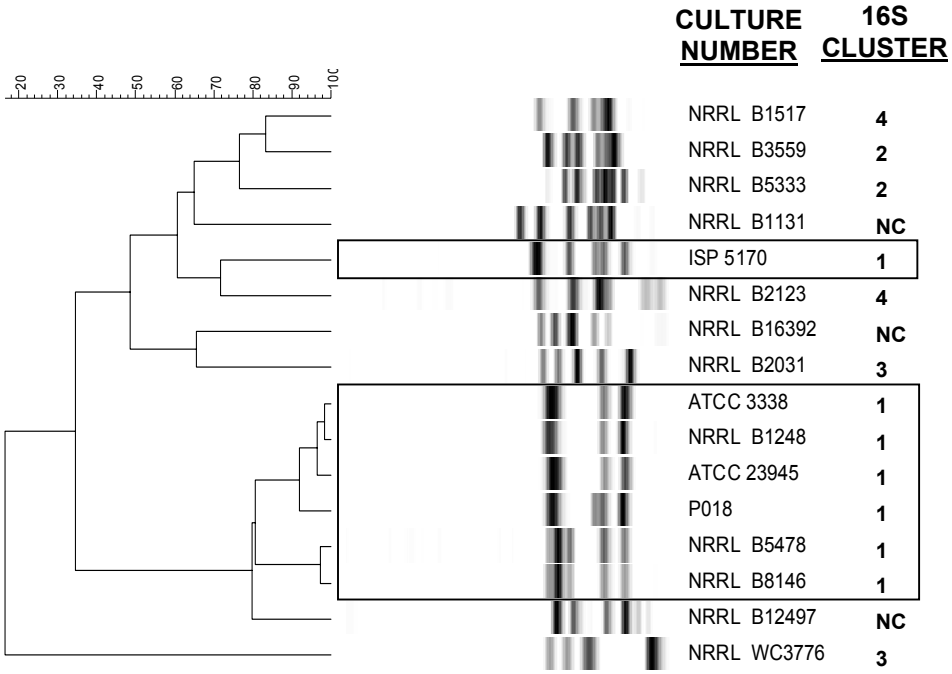


Figure 3.4. Ribotype comparison of *S. phaeochromogenes* and closely related strains.

Ribotyping was performed using the RiboPrinter Microbial Characterization System (Dupont Qualicon, Wilmington, DE) and cluster analysis of Riboprints was performed using BioNumerics software (Applied Maths, Austin, TX). Boxes highlight the strains corresponding to 16S rRNA gene cluster 1 (Figure 3.2). NC = not clustered.

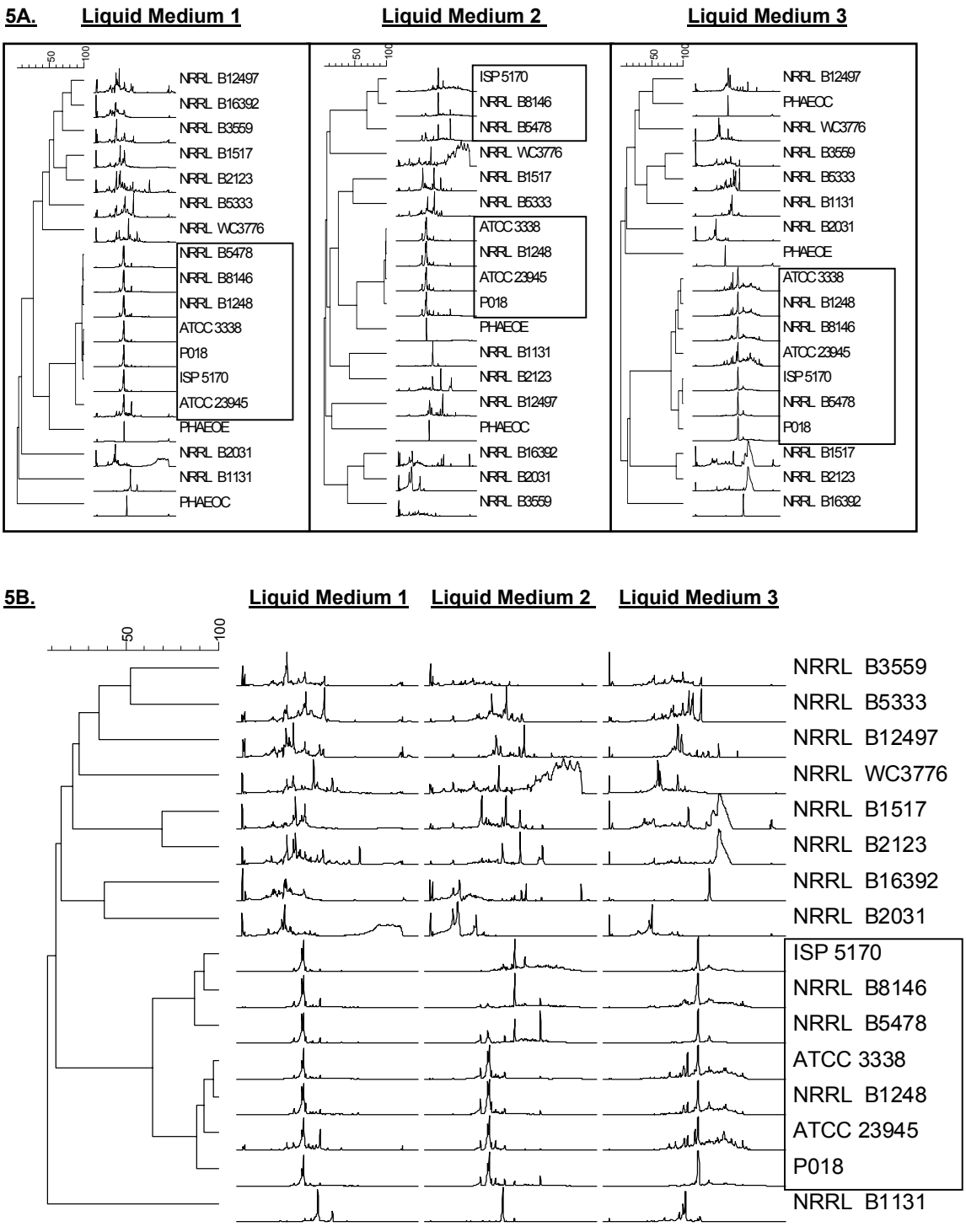


Figure 3.5. Metabolite profiles of *S. phaeochromogenes* and closely related strains. in three liquid media.

Cluster analysis of HPLC chromatograms was performed using BioNumerics software (Applied Maths, Austin, TX) using the unweighted pair group method using arithmetic averages (UPGMA). Boxes highlight the strains corresponding to 16S rRNA gene cluster 1 (Figure 3.2). 5A: Separate comparisons of metabolite profiles in individual media. 5B: Composite comparison of all three metabolite profiles. For composite comparison, each data set was weighted equally.

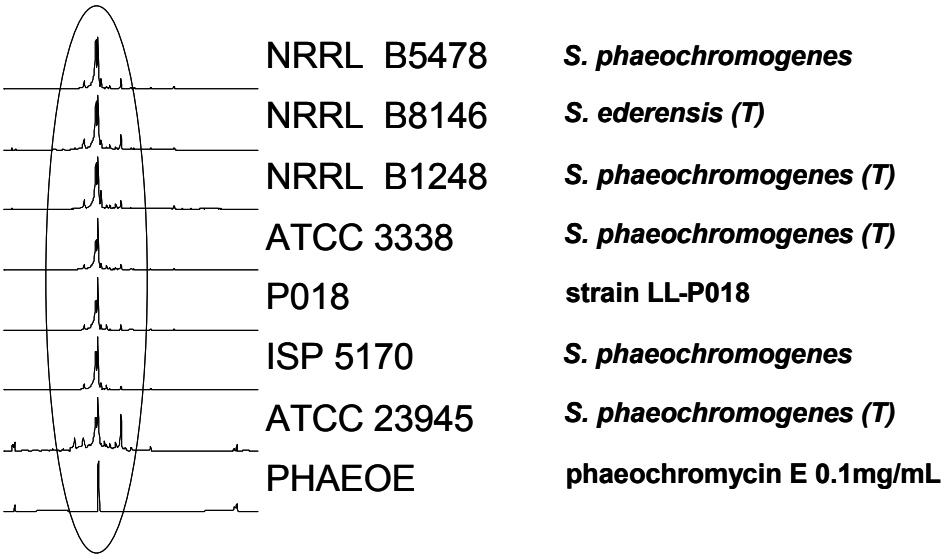


Figure 3.6. HPLC analysis of phaeochromycin production by *S. phaeochromogenes* and closely related strains.

All of the strains in 16S rRNA gene cluster 1 produce the phaeochromycins (above) and alnumycin (not shown). PHAEOE = phaeochromycin E, 0.1 mg/mL; (T) = type strain.

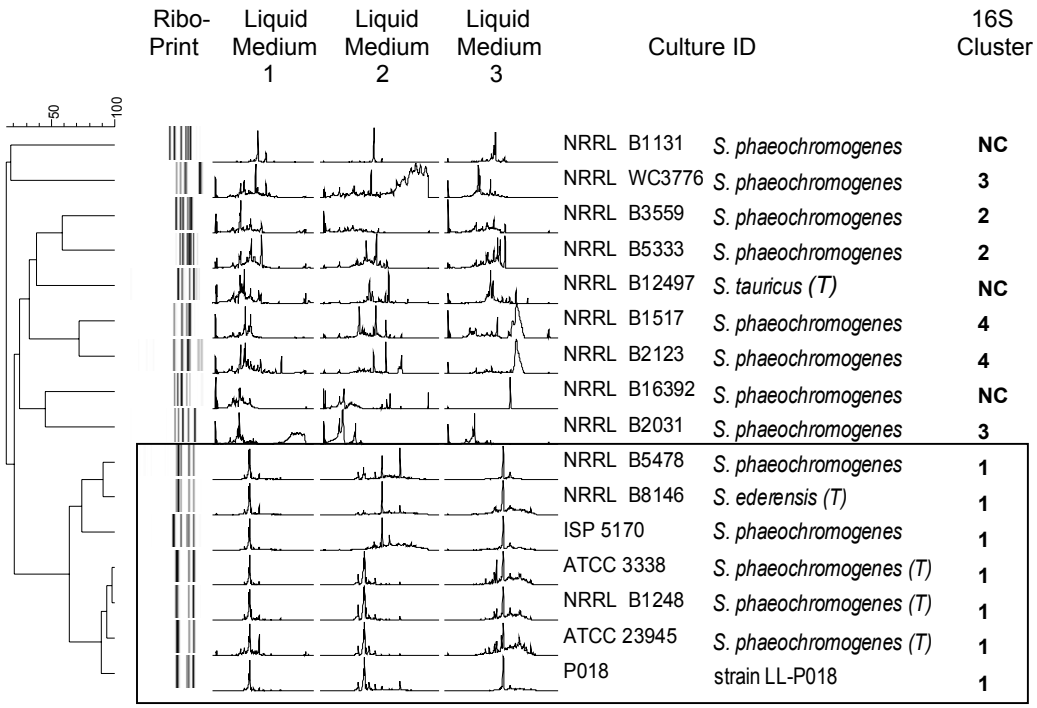


Figure 3.7. Combined cluster analysis of *S. phaeochromogenes* and closely related strains by ribotype and metabolite profiles.

Cluster analysis of Riboprints and HPLC chromatograms was performed using BioNumerics software (Applied Maths, Austin, TX) using the unweighted pair group method using arithmetic averages (UPGMA). For composite comparison, each data set was weighted equally. Boxes highlight the strains corresponding to 16S rRNA gene cluster 1 (Figure 3.2). NC = not clustered; (T) = type strain.

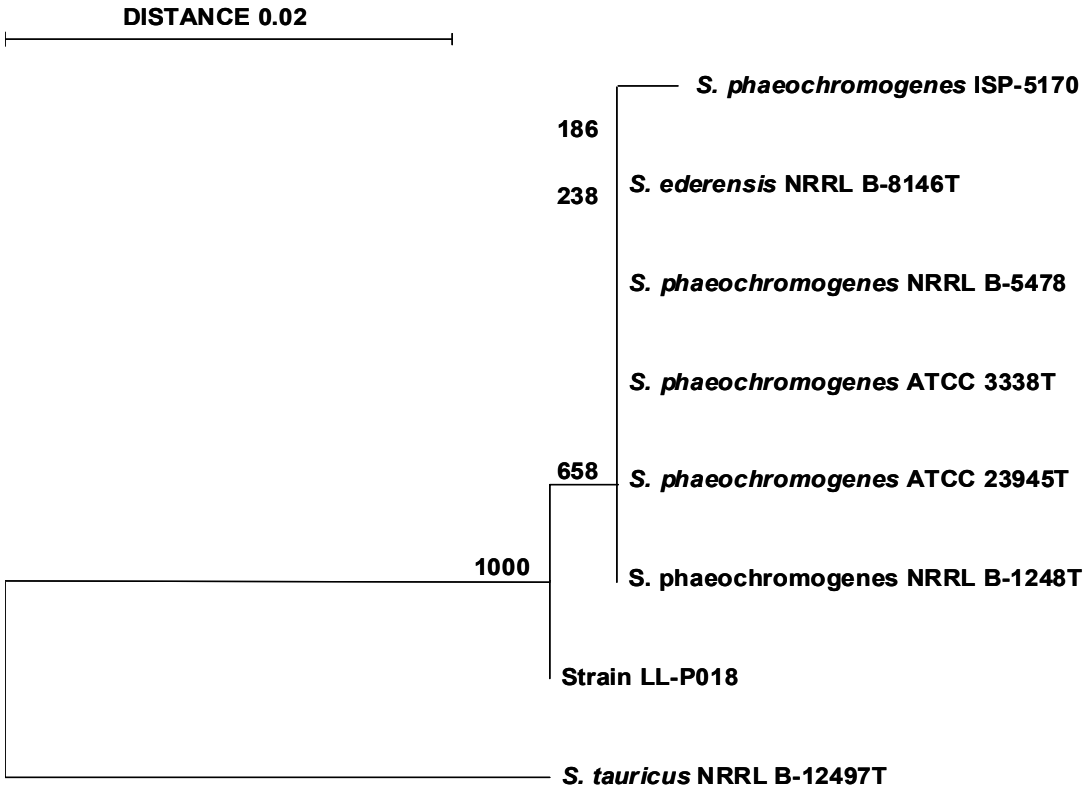


Figure 3.8. Comparison of *S. phaeochromogenes* and closely related strains by *rpoB* sequence.

Ribosomal polymerase β-subunit (*rpoB*) gene sequences (354bp) were aligned using ClustalX (Thompson *et al.*, 1997). The tree was constructed with TREECON software using the neighbor-joining method with 1000 bootstrap replicates (Jukes & Cantor, 1969; Van de Peer & De Wachter, 1994). T = type strain.

Culture Number	Presumptive Identification	Source
LL-P018	*Streptomyces* strain LL-P018	Wyeth Research, Natural Products Culture Collection, Pearl River, NY.
ATCC 3338^T	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	American Type Culture Collection, Manassas, VA.
ATCC 23945^T	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	American Type Culture Collection, Manassas, VA.
NRRL B-1248^T	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-1131	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-1517	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-2031	*Streptomyces phaeochromogenes* subspecies *chloromyceticus*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-2123	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-3559	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-5333	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-5478	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-16392	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
ISP-5170	*Streptomyces phaeochromogenes* subspecies *phaeochromogenes*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
WC-3776	*Streptomyces phaeochromogenus* subspecies *chloromyceticus*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-8146^T	*Streptomyces ederensis*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.
NRRL B-12497^T	*Streptomyces tauricus*	Agricultural Research Service (ARS) Culture Collection, USDA, Peoria, IL.

Table 3.1. *S. phaeochromogenes* and closely related strains. (T) = type strain.

Carbohydrate utilization				
Carbon source	***Streptomyces* strain LL-P018**	***Streptomyces phaeochromogenes* ATCC 3338^T**	***Streptomyces ederensis* NRRL B-8146^T**	***Streptomyces tauricus* NRRL B-12497^T**
Glucose	+	+	+	+
Arabinose	+	+	+	+
Sucrose	+	+	+	+
Xylose	+	+	+	+
Inositol	+	+	+	+
Mannitol	+	+	+	+
Fructose	+	+	+	+
Rhamnose	+	+	+	+
Raffinose	+	+	+	+
Cellulose	+/-	-	+/-	-
No carbon	-	-	-	-

Table 3.2. Carbohydrate utilization by *S. phaeochromogenes* and closely related species.

Utilization of various carbohydrates was compared according to methods defined by the International Streptomyces Project (Shirling & Gottlieb, 1966). "+" = good growth, "+/-" = weak growth, "-" = no growth.

ISP sporulation agar morphology and pigmentation				
ISP Medium	***Streptomyces* strain LL-P018**	***Streptomyces phaeochromogenes* ATCC 3338^T**	***Streptomyces ederensis* NRRL B-8146^T**	***Streptomyces tauricus* NRRL B-12497^T**
ISP2	sub myc:	sub myc:	sub myc:	sub myc:
(yeast extract-	dark brown	brown	dark brown	pink/tan
malt extract agar)	rev myc:	rev myc:	rev myc:	rev myc:
	dark brown	brown	dark brown	pink/tan
	aerial myc:	aerial myc:	aerial myc:	aerial myc:
	white	white	white	none
	spore mass:	spore mass:	spore mass:	spore mass:
	grey	grey	grey	none
	pigment:	pigment:	pigment:	pigment:
	brown	brown	brown	none
ISP3	sub myc:	sub myc:	sub myc:	sub myc:
(oatmeal agar)	brown	tan	brown	pink/brown
	rev myc:	rev myc:	rev myc:	rev myc:
	brown	tan	brown	pink/brown
	aerial myc:	aerial myc:	aerial myc:	aerial myc:
	white/pink	white/grey	white/pink	white/pink
	spore mass:	spore mass:	spore mass:	spore mass:
	white/pink	white/grey	white/pink	white/pink
	pigment:	pigment:	pigment:	pigment:
	brown	none	trace-brown	none
ISP4	sub myc:	sub myc:	sub myc:	sub myc:
(inorganic salts-	brown	tan	brown	pink
starch agar)	rev myc:	rev myc:	rev myc:	rev myc:
	brown	tan	brown	pink
	aerial myc:	aerial myc:	aerial myc:	aerial myc:
	white/pink	white	white/pink	white/pink
	spore mass:	spore mass:	spore mass:	spore mass:
	white/pink	white	white/pink	white/pink
	pigment:	pigment:	pigment:	pigment:
	none	none	none	none
ISP5	sub myc:	sub myc:	sub myc:	sub myc:
(glycerol-asparagine	dark brown	brown	dark brown	pink/tan
agar)	rev myc:	rev myc:	rev myc:	rev myc:
	dark brown	brown	dark brown	pink/tan
	aerial myc:	aerial myc:	aerial myc:	aerial myc:
	grey	grey	grey	white/pink
	spore mass:	spore mass:	spore mass:	spore mass:
	grey	grey	grey	none
	pigment:	pigment:	pigment:	pigment:
	tan	tan	tan	none

Table 3.3. Morphological comparison of *S. phaeochromogenes* and closely related species.

Growth, sporulation, and pigmentation were observed according to methods defined by the International Streptomyces Project (Shirling & Gottlieb, 1966).

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Anderson, A. S. & Wellington, E. M. H. (2001). The taxonomy of *Streptomyces* and related genera. *Int J Syst Evol Microbiol* **51**, 797-814.

Anzai, Y., Okuda, T. & Watanabe, J. (1997). Application of the random amplified polymorphic DNA using the polymerase chain reaction for efficient elimination of duplicate strains in microbial screening. II. Actinomycetes. *J Antibiot* **50**, C-1.

Baltz, R. H. (1997). Lipopeptide antibiotics produced by *Streptomyces roseosporus* and *Streptomyces fradiae*. In *Drugs and the Pharmaceutical Sciences*, pp. 415-435. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Baltz, R. H. (2006). Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J Ind Microbiol Biotechnol* **33**, 507-513.

Beijerinck, M. W. (1900). Ueber Chinonbildung durch *Streptothrix chromogena* und Lebensweise dieses Mikroben. *Centr Bakteriol Parasitenk Abt II* **6**, 2-12. Cited in: Waksman (1959) and Waksman and Curtis (1916).

Bruce, J. (1996). Automated system rapidly identifies and characterizes microorganisms in food. *Food Tech* **50**, 77-81.

Challis, G. L. & Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci USA* **100 (Suppl 2)**, 14555-14561.

Chater, K. F. & Bibb, M. J. (1997). Regulation of bacterial antibiotic production. In *Biotechnology (2nd Edition)*, pp. 57-105. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Clardy, J., Fischbach, M. A. & Walsh, C. T. (2006). New antibiotics from bacterial natural products. *Nat Biotechnol* **24**, 1541-1550.

Conn, H. J. (1917). Soil Flora Studies. V. Actinomycetes in Soil. *NY Agr Exp Sta Tech Bull* **60**, 1-16.

Graziani, E. I., Ritacco, F. V., Bernan, V. S. & Telliez, J. B. (2005). Phaeochromycins A-E, anti-inflammatory polyketides isolated from the soil actinomycete *Streptomyces phaeochromogenes* LL-P018. *J Nat Prod* **68**, 1262-1265.

Hansen, J. N. (1997). Nisin and related antimicrobial peptides. In *Drugs and the Pharmaceutical Sciences*, pp. 437-470. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Jack, R., Gotz, F. & Jung, G. (1997). Lantibiotics. In *Biotechnology (2nd Edition)*, pp. 323-368. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mammalian protein metabolism*, pp. 21-132. Ed. H. H. Munro. Academic Press, New York.

Kim, B. J., Kim, C. J., Chun, J., Koh, Y. H., Lee, S. H., Hyun, J. W., Cha, C. Y. & Kook, Y. H. (2004). Phylogenetic analysis of the genera *Streptomyces* and *Kitasatospora* based on partial RNA polymerase beta-subunit gene (*rpoB*) sequences. *Int J Syst Evol Microbiol* **54**, 593-598.

Kotlyarov, A., Neininger, A., Schubert, C., Eckert, R., Birchmeier, C., Volk, H. D. & Gaestel, M. (1999). MAPKAP kinase 2 is essential for LPS-induced TNF-alpha biosynthesis. *Nat Cell Biol* **1**, 94-97.

Labeda, D. (2006). *rpoB* Gene Amplification and Sequencing (after Kim et al., 2004, IJSEM 54:593-598). United States Department of Agriculture, Peoria, IL. Personal communication.

Lanoot, B., Vancanneyt, M., Hoste, B., Vandemeulebroecke, K., Cnockaert, M. C., Dawyndt, P., Liu, Z., Huang, Y. & Swings, J. (2005). Grouping of streptomycetes using 16S-ITS RFLP fingerprinting. *Res Microbiol* **156**, 755-762.

Mankelow, D. P. & Neilan, B. A. (2000). Non-ribosomal peptide antibiotics. *Exp Op Therap Pat* **10**, 1583-1591.

Nicas, T. I. & Cooper, R. D. G. (1997). Vancomycin and other glycopeptides. In *Drugs and the Pharmaceutical Sciences*, pp. 363-392. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Pearson, K. (1926). On the coefficient of racial likeness. *Biometrika* **18**, 105-117.

Pfaller, M. A., Wendt, C., Hollis, R. J., Wenzel, R. P., Fritschel, S. J., Neubauer, J. J. & Herwaldt, L. A. (1996). Comparative evaluation of an automated ribotyping system versus pulsed-field gel electrophoresis for epidemiological typing of clinical isolates of *Escherichia coli* and *Pseudomonas aeruginosa* from patients with recurrent gram-negative bacteremia. *Diagn Microbiol Infect Dis* **25**, 1-8.

Piepersberg, W. & Distler, J. (1997). Aminoglycosides and sugar components in other secondary metabolites. In *Biotechnology (2nd Edition)*, pp. 397-488. Ed. H. Kleinkauf & H. Von Doehren. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Reysenbach, A. L., Wickham, G. S. & Pace, N. R. (1994). Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl Environ Microbiol* **60**, 2113-2119.

Ritacco, F. V., Haltli, B., Janso, J. E., Greenstein, M. & Bernan, V. S. (2003). Dereplication of *Streptomyces* soil isolates and detection of specific biosynthetic genes using an automated ribotyping instrument. *J Ind Microbiol Biotechnol* **30**, 472-479.

Rullmann, W. (1895). Chemisch-bakteriologische Untersuchungen von Zwischendeckenfüllungen, mit besonderer Berücksichtigung von *Cladothrix odorifera.* Munich, Germany: Universität München. Cited in: Waksman (1959) and Waksman and Curtis (1916).

Schatz, A., Bugie, E. & Waksman, S. A. (1944). Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc Soc Exp Biol Med* **55**, 66-69.

Sehgal, S. N., Baker, H. & Vezina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. *J Antibiot (Tokyo)* **28**, 727-732.

Shirling, E. B. & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313-340.

Strohl, W. R. (1997). Industrial antibiotics: today and the future. In *Drugs and the Pharmaceutical Sciences*, pp. 1-47. Ed. W. R. Strohl. Marcel Dekker, Inc., New York.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876-4882.

Ueda, K., Seki, T., Kudo, T., Yoshida, T. & Kataoka, M. (1999). Two distinct mechanisms cause heterogeneity of 16S rRNA. *J Bacteriol* **181**, 78-82.

Van de Peer, Y. & De Wachter, R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci* **10**, 569-570.

Vauterin, L. & Vauterin, P. (2006). Integrated Databasing and Analysis. In *Molecular Identification, Systematics, and Population Structure of Prokaryotes*, pp. 141-217. Ed. E. Stackebrandt. Springer-Verlag, Heidelberg, Germany.

Waksman, S. A. (1959). *The Actinomycetes, Vol.1: Nature, Occurrence, and Activities*.The Williams & Wilkins Co., Baltimore, MD.

Waksman, S. A. & Curtis, R. E. (1916). The actinomyces of the soil. *Soil Sci* **1**, 99-134.

Waksman, S. A. & Woodruff, H. B. (1942). Streptothricin, a new selective bacteriostatic and bactericidal agent, particularly active against gram-negative bacteria. *Proc Soc Exp Biol Med* **49**, 207-210.

Walsh, C. T. (2003). *Antibiotics: actions, origins, resistance*.ASM Press, Washington, D.C.

Walsh, C. T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805-1810.

Wink, J. (2002). *The Actinomycetales: an order in the class of Actinobacteria important to the pharmaceutical industry- electronic manual*.CD-Rom Provided by Dr. Joachim Wink, Aventis Pharma Deutschland GmbH.

Chapter 4

Biosynthetic genes involved in both phaeochromycin and alnumycin biosynthesis.

Frank V. Ritacco[1], Edmund I. Graziani[1], Brad Haltli[1], and Douglas E. Eveleigh[2].

[(1)] Wyeth Research, Pearl River, NY, and [(2)] School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ.

Abstract

The phaeochromycins are novel anti-inflammatory polyketides produced by *Streptomyces phaeochromogenes* strain LL-P018. Their biosynthesis was examined through biochemical and genetic analysis. Growth of strain LL-P018 in a variety of liquid media yielded phaeochromycins and an additional pigmented compound. The latter was purified by reverse-phase preparative HPLC and then identified as alnumycin. In order to identify the genes for phaeochromycin biosynthesis, a cosmid library was screened for the presence of polyketide biosynthetic genes. A 4.1kb contiguous DNA sequence was identified which contains 5 open reading frames that encode type II polyketide synthase genes. A knockout construct based on this sequence, when introduced into strain LL-P018, yielded mutants unable to produce either the phaeochromycins or alnumycin, indicating that these polyketides are products of a single biosynthetic pathway.

Introduction

Streptomyces phaeochromogenes strain LL-P018 produces the phaeochromycins (Figure 4.1), a new family of biologically active polyketides (Graziani *et al.*, 2005; Ritacco & Eveleigh, 2007). Two members of this family, phaeochromycins A and C, are inhibitors of MK-2 kinase, a promising target in the search for new anti-inflammatory therapies (Graziani *et al.*, 2005). MK-2 (MAP kinase-activated kinase 2) is an essential component in the production of tumor necrosis factor (TNF-α) and other inflammatory cytokines leading to the inflammatory response (Kotlyarov *et al.*, 1999). Because elimination of MK-2 activity can dramatically reduce TNF-α production, small molecule inhibitors of MK-2 are of interest as potential anti-inflammatory drug candidates.

The phaeochromycins (Figure 4.1) are analogous to partially cyclized shunt products of the actinorhodin and enterocin biosynthetic pathways (Graziani *et al.*, 2005; Kalaitzis & Moore, 2004; McDaniel *et al.*, 1994a, b; Xiang *et al.*, 2002). This similarity suggests that the phaeochromycins themselves may be partially cyclized intermediates or shunt metabolites derived from a biosynthetic pathway for which the fully cyclized end product has not yet been identified. In order to understand the biosynthesis of these newly discovered polyketides, strain LL-P018 was grown in several liquid culture media and evaluated for the production of phaeochromycins and new metabolites. Metabolites of interest were purified by preparative HPLC and structural information was obtained by NMR analysis.

In order to identify the phaeochromycin biosynthetic gene cluster, genomic DNA from strain LL-P018 was examined for the presence of type II polyketide synthase (PKS) genes by screening using degenerate PCR (polymerase chain reaction). A PCR product representing a putative ketosynthase gene was then used to probe an LL-P018 genomic cosmid library. By this method, a portion of a type II PKS pathway from strain LL-P018 has been identified. Potential functions of these genes have been assigned based on BLAST analysis (Altschul *et al.*, 1997), and a knockout construct was developed and used to confirm the role of these genes in the biosynthesis of the phaeochromycins and other metabolites.

Materials and Methods

Culture conditions

Strain LL-P018 was maintained at 28°C on solid medium containing (g/L): D-glucose (dextrose) (10), yeast extract (Difco/Becton-Dickinson, Franklin Lakes, NJ) (5), agar (15), soluble starch (Difco/Becton-Dickinson, Franklin Lakes, NJ) (20), NZ-amine A (Sheffield/Kerry Bioscience, Norwich, NY) (5), and calcium carbonate (1). For production of the phaeochromycins and alnumycin, one loopful (10μL) of culture from agar plates was inoculated into 10mL of seed medium containing (g/L): D-glucose (10), soluble starch (Difco/Becton-Dickinson, Franklin Lakes, NJ) (20), yeast extract (Difco/Becton-Dickinson, Franklin Lakes, NJ) (5), NZ-amine A (Sheffield/Kerry Bioscience, Norwich, NY) (5), and calcium carbonate (1). These seed cultures were

incubated for 3 days at 28°C with shaking at 200 rpm and 2 inch throw. One liter of production medium was inoculated with 20mL of seed culture (2% inoculum) and the culture incubated for 7 days at 28°C with shaking at 200 rpm and 2 inch throw. Phaeochromycin production medium contained (g/L): magnesium sulfate 7-hydrate (0.5), potassium chloride (0.5), dibasic potassium phosphate (2.5), sodium chloride (5), glycerol (10), and soy peptone (5). Alnumycin production medium contained (g/L): D-glucose (dextrose) (1), D-mannitol (10), Bacto-peptone (Difco/Becton-Dickinson, Franklin Lakes, NJ) (1), soluble starch (Difco/Becton-Dickinson, Franklin Lakes, NJ) (15), magnesium sulfate 7-hydrate (0.1), monobasic potassium phosphate (0.5), and calcium carbonate (1).

Analysis of culture broths for polyketides

Whole broth was extracted in ethyl acetate (1:1) and centrifuged to separate phases. The organic phase was dried and the product resolubilized in methanol to a 10-fold final concentration. Reverse phase HPLC of extracts was performed using an Agilent model 1100 liquid chromatograph with photodiode array detection, coupled to an Agilent LC-MSD ion trap mass spectrometer fitted with an electrospray ionization source. Components in the extracts were resolved by reverse phase chromatography using a Zorbax Eclipse XDB-C18 column (3.0 x 75mm, 3.5µm particle size) with a mobile phase of 0.025% formic acid in water (solvent A) and 0.025% formic acid in acetonitrile (solvent B). For elution, a linear gradient from 5% A to 95% B in 15 minutes was used, holding at 95% acetonitrile for 10 minutes, with a flow rate of 0.8 mL/minute. A split diverted 25% of the flow into the mass spectrometer. UV detection was performed using

a photodiode array with a scan range of 200-600 nm. The MS electrospray was performed in positive mode with a scan range of 200-2,000 m/z.

Purification of alnumycin

Strain LL-P018 was cultured in one liter of alnumycin production medium as described above. Whole broth (1L) was extracted three times with 300mL of ethyl acetate in a 1L separatory funnel. The organic layer (900mL) was collected and dried to completion in a rotary evaporator. The dried material was resolubilized in methanol and coated onto silica (MPsilica 32-63, 60Å, MP Biomedicals, Germany). The charged silica was loaded on top of a 100mL bed of fresh hexane-washed silica in a Buchner funnel and washed with three bed volumes (300mL) of hexane. Using vacuum liquid chromatography (VLC), alnumycin (Figure 4.2) was partially eluted from the silica with 300mL of 50% ethyl acetate/hexane, and the remainder was eluted with 600mL of 100% ethyl acetate. These fractions were dried to completion in a rotary evaporator yielding 500mg of dried crude material. The crude material was resolubilized in 8mL of 50% acetonitrile, 25% water, 25% methylsulfoxide (DMSO) and injected onto a preparatory HPLC column (YMC Pack ODS S-10P 50 x 250mm, 120Å). Alnumycin was eluted using a gradient of 50% to 95% acetonitrile in water at a flow rate of 20mL/minute for 125 minutes (20mL fraction collected per minute). Fractions 65-71, which were highly enriched for alnumycin, were pooled (140mL total) and dried to completion in a rotary evaporator. Dried material was resolubilized in deuterated chloroform ($CDCl_3$) and examined by proton (^{1}H) and carbon (^{13}C) NMR. NMR spectra were recorded on a Bruker Avance DPX-400 spectrometer using a 3 mm broadband detect probe.

Preparation of a type II PKS DNA probe for the phaeochromycin biosynthetic gene cluster

Genomic DNA was isolated from *Streptomyces* strain LL-P018 according to Kieser *et al*. (2000). A pair of degenerate primers based on highly conserved type II polyketide synthase (PKS) gene sequences in *Streptomyces* (Metsa-Ketela *et al.*, 1999) was used to amplify α-ketosynthase (KSα) gene fragments from strain LL-P018. The PCR was performed using JumpStart DNA polymerase (Sigma, St. Louis, MO) according to the manufacturer's instructions. Primers were used at a final concentration of 10μM. A BioMetra T Gradient thermocycler was programmed for PCR reactions as follows: one cycle of denaturation at 95ºC for 2 minutes, followed by 35 cycles of denaturation at 95ºC for 1 minute, annealing at 65ºC for 45 seconds, and extension at 72ºC for 90 seconds, ending with one extension cycle at 72ºC for 5 minutes and a pause at 4ºC. A 600bp PCR product was excised from an agarose gel, the DNA was purified using the Zymo Gel DNA Recovery Kit (Zymo Research, Orange, CA) and then cloned using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The plasmid insert was directly sequenced using an ABI 3700 sequencer with the ABI Prism DNA sequencing kit and Big Dye terminators version 3.0 (Applied Biosystems, Foster City, CA). Sequencing was performed on the plasmid insert using the M13 forward and reverse primers, and the sequence was compared to the GenBank database by nucleotide BLAST search (Altschul *et al.*, 1997). Based on homology to known sequences in GenBank, the 600bp insert was determined to be a portion of a KSα gene from a type II PKS gene cluster as predicted (Figure 4.3, Table 4.1). This DNA

fragment was labeled with ^{32}P -dCTP by random priming using the Megaprime kit (Amersham/GE Healthcare Biosciences, Piscataway, NJ) and used to probe a cosmid library.

Cosmid screening

Genomic DNA from strain LL-P018 was sheared by fourteen passes through a 1mL 27G½ syringe (Becton-Dickinson, Franklin Lakes, NJ), yielding 20µg of sheared DNA fragments approximately 40kb in size. A cosmid library was then generated using the pWEB cosmid cloning kit (Epicentre Biotechnologies, Madison, WI) according to the manufacturer's instructions. Three hundred sixty-two cosmid clones were obtained, representing approximately 31% coverage of the LL-P018 genome (assuming a genome size of 8-9Mb). Cosmid clones were screened with the DNA probe described above using standard protocols for colony hybridization (Sambrook *et al.*, 1989), and two positive clones were identified. Cosmid DNA was isolated from these positive clones using the Qiagen Large-Construct Kit (Qiagen, Valencia, CA). Cosmid inserts were partially sequenced using the KSα PCR primers described above (Metsa-Ketela *et al.*, 1999) to obtain sequence for the regions flanking the known KSα probe sequence, and the M13 forward and T7 reverse primers were used to obtain end sequence data for the cosmid insert. Sequence data from both cosmids was assembled using Sequencher software (Gene Codes Corporation, Ann Arbor, MI), and open reading frames (ORFs) were identified using FramePlot (Ishikawa & Hotta, 1999).

Disruption of the phaeochromycin biosynthetic pathway

Based on sequence data obtained from two cosmids, a 4.1kb contiguous DNA sequence was identified with homology to type II minimal PKS genes (see Results). Based on this sequence, PCR primers were designed and used to amplify a 2.1kb fragment containing the entire β –ketosynthase (KSβ) gene along with 390 bp of upstream flanking sequence and 476 bp of downstream flanking sequence. The primers used were PE-33607F (5'-GGCGCAGAGGTAGACGG-3'), corresponding to nucleotide positions 1396-1411 of the 4.1kb sequence and PE-33607R (5'- CTCACTGGACCAGGCG -3'), corresponding to nucleotide positions 3532-3548 (complementary strand). PCR reactions were performed as described above for amplification of the KSα gene.

The resultant 2.1kb PCR product was cloned into pCR-BluntII-TOPO (Invitrogen, Carlsbad, CA) creating plasmid pFRPH01. A unique NruI site in the center of the KSβ gene, with approximately 1kb of sequence flanking each side, was used for introduction of a resistance marker. A cassette containing an origin of transfer (OriT) along with an apramycin resistance (AprR) gene was amplified from plasmid pJV177 (provided by Nathan Magarvey, unpublished) by PCR. Plasmid pFRPH01 was digested with NruI, and the OriT-AprR cassette was blunt-end ligated into the NruI site, creating plasmid pFRPH01KO (Figure 4.4). The presence of the insert was confirmed by digestion of the plasmid with DraI, yielding a 7.4 kb fragment. The orientation of the insert was confirmed by digestion of the plasmid with XbaI and MluI, yielding two fragments of approximately 2.7kb and 5kb. These results indicated that the insert was in reverse orientation relative to the biosynthetic genes.

Plasmid pFRPH01KO was transformed into the methylation deficient *Escherichia coli* donor strain ET12567 (pUZ8002), and then subsequently transferred to strain LL-P018 *via* conjugation with *Escherichia coli* using standard methods (Flett *et al.*, 1997; Kieser *et al.*, 2000). Exconjugants were isolated on the agar medium described above, supplemented with 50μg/mL apramycin. Because this plasmid lacks a streptomycete origin of replication and is therefore unable to replicate in *Streptomyces*, apramycin-resistant exconjugants were presumed to have plasmid pFRPH01KO integrated into the native gene cluster *via* homologous recombination (Figure 4.5). Exconjugants were screened for the ability to produce the phaeochromycins and alnumycin by growth in liquid culture media, followed by HPLC-MS analysis as described above.

A PCR screen was devised in order to confirm the presence of the gene disruption in apramycin resistant colonies. Primers based on nucleotide positions 2196-2213 (5'-GGAGATCCGCAAGCTGTG-3') and 2616-2636 (complementary strand; 5'-GAGAAGGTGGCCGCGTATCCG-3') of the 4.1kb sequence were designed to amplify a fragment of the KSβ gene. Amplification of the wild-type gene with these primers was expected to yield a 440bp product, whereas amplification of the disrupted gene was predicted to yield a 2.1 kb fragment (Appendix 8). Clones which have integrated the disrupted gene construct *via* double-crossover would be identified by amplification of the 2.1kb fragment and the absence of the 440bp fragment. A single crossover, however, would only introduce the disrupted copy of the gene without eliminating the native copy, and therefore both amplification products would be detected. A second set of primers was

designed to amplify plasmid backbone DNA from plasmid pFRPH01KO. The primers used were (forward) 5'-GACCCCGTAGAAAAGATC-3' and (reverse) 5'-CATGTGAGCAAAAGGCCA-3' corresponding to nucleotide positions 2721-2738 and 3380-3398 of the pCR-BluntII-TOPO plasmid sequence. Clones which have integrated the disrupted gene construct *via* double-crossover would be identified by the lack of amplification of any plasmid DNA, whereas clones which are only disrupted *via* single crossover would still retain the plasmid backbone, and therefore amplification of a 677-bp fragment would be expected.

Results

Production of phaeochromycins and related metabolites by strain LL-P018

In phaeochromycin production media, strain LL-P018 produced the phaeochromycins (Figures 4.1 and 4.6). Clear separation of phaeochromycins A, B, and C (elution times 5.5 minutes, 5.4 minutes, and 5.7 minutes respectively) was obtained. Production of phaeochromycins was not observed when strain LL-P018 was grown in a second liquid medium (later named alnumycin production medium). However, HPLC-MS analysis of culture extracts from this medium revealed a major metabolite peak that did not correspond with any of the known phaeochromycins by either retention time or molecular mass (Figure 4.6). This metabolite eluted at approximately 9.0 minutes and had a mass of 417.2 in the positive scan.

Isolation and drying of this metabolite yielded a dark red powder, which was preliminarily identified as alnumycin (Figure 4.2) by mass and UV spectrometry (Bieber *et al.*, 1998). Alnumycin is also referred as K1115B$_1$ (Naruse *et al.*, 1998). Carbon and proton NMR spectra were in agreement with previously reported values for alnumycin and K1115B$_1$ (Bieber *et al.*, 1998; Naruse *et al.*, 1998).

Identification of putative phaeochromycin biosynthetic genes

A 600bp PCR product was amplified from the genome of strain LL-P018 using degenerate primers based on conserved sequence from type II ketosynthase genes (Metsa-Ketela *et al.*, 1999). Sequencing revealed this DNA fragment to be homologous to KSα gene sequences from type II PKS gene clusters, including those for actinorhodin and frenolicin (Bibb *et al.*, 1994; Fernandez-Moreno *et al.*, 1992; Malpartida & Hopwood, 1984), suggesting that it may be a component of the phaeochromycin biosynthetic machinery.

This KSα gene fragment was radiolabeled and used to probe a cosmid library representing the genome of strain LL-P018. Because the cosmid titer was low, only 362 cosmid clones were screened, representing approximately 31% of the genome. However, two positive clones were identified which hybridized with the radiolabeled probe. Isolation and partial sequencing of cosmid DNA from these two clones revealed a 4.1kb contiguous DNA sequence with homology to type II minimal PKS genes. FRAME analysis and comparison of translated open reading frame (ORF) sequences to the GenBank protein database identified 5 genes corresponding to a type II PKS acyl

transferase (AT), ketosynthase α- and β-subunits (KSα, KSβ), an acyl carrier protein (ACP) and a ketoreductase (KR) (Figure 4.3, Table 4.1). All of the deduced protein sequences were homologous to enzymes from known type II PKS biosynthetic pathways, including those for frenolicin and R1128 (Bibb *et al.*, 1994; Marti *et al.*, 2000).

Elimination of phaeochromycin and alnumycin biosynthesis by gene disruption

Plasmid construct pFRPH01KO was designed to disrupt the function of the putative KSβ gene *via* homologous recombination and insertion of an apramycin resistance marker (Figures 4.4 and 4.5). Transfer of this construct to strain LL-P018 *via* intergeneric conjugation with *E. coli* ET12567/pUZ8002/pFRPH01KO yielded approximately 30 apramycin-resistant colonies. Selected isolates (mutants XC20, XC21, and XC22) grown in liquid culture media were unable to produce either the phaeochromycins or alnumycin (Figure 4.6).

PCR screening of mutants unable to produce the phaeochromycins or alnumycin in liquid culture indicated that the KSβ was disrupted *via* single-crossover homologous recombination. Amplification of the KSβ gene from mutants XC20, XC21, and XC22 yielded two PCR products, approximately 2.1 kb and 440bp in size (Appendix 8). These results matched the predicted sizes of products expected to result from the integration of the disrupted gene *via* a single-crossover at one of the homology regions flanking the resistance marker (Figure 4.5). pCR-BluntII-TOPO plasmid backbone DNA was also successfully amplified from these mutants (Appendix 9). These results, combined with the elimination of phaeochromycin and alnumycin production in these mutants, suggest

that this pathway was successfully disrupted *via* single-crossover at the KSβ gene, and that this gene cluster is responsible for the biosynthesis of the phaeochromycins and alnumycin in strain LL-P018.

Discussion

The phaeochromycins (Figure 4.1) are aromatic polyketides, and their biosynthesis is likely catalyzed by a type II polyketide synthase (PKS) as described for this class of molecules (Hopwood, 1997; Kalaitzis & Moore, 2004; Keatinge-Clay *et al.*, 2004). Structurally, the phaeochromycins closely resemble shunt products from the actinorhodin and enterocin biosynthetic pathways, differing in the presence of a propyl side chain possibly derived from a butyrate starter unit (Graziani *et al.*, 2005). Phaeochromycins A and B are analogous to the actinorhodin biosynthetic shunt products dehydromutactin and mutactin (Kalaitzis & Moore, 2004; Zhang *et al.*, 1990) respectively, as well as the enterocin shunt products wailupemycin G and F (Xiang *et al.*, 2002, 2004). These particular mutactins and wailupemycins are naturally produced at very low levels by their respective wild-type producing organisms. However, significant levels are produced in mutants lacking specific biosynthetic genes involved in cyclization and aromatization of the linear polyketide chain (Kalaitzis & Moore, 2004; Piel *et al.*, 2000; Xiang *et al.*, 2002). Similarly, phaeochromycin C is analogous to the actinorhodin shunt product SEK34b (McDaniel *et al.*, 1994b) and phaeochromycin E is analogous to the actinorhodin shunt product BSM1 (Kalaitzis & Moore, 2004). SEK34b and BSM1 have not been described as metabolites of wild-type organisms. SEK34b is produced by a

mutant of the actinorhodin producing strain *S. coelicolor* lacking certain biosynthetic genes (McDaniel *et al.*, 1994b), and BSM1 is produced by expression of an incomplete set of actinorhodin biosynthetic genes in the heterologous host organism *S. lividans* K4-114 (Kalaitzis & Moore, 2004). In the case of BSM1, this has led to the speculation that this metabolite may not be a true biosynthetic product, but instead a catabolic product resulting from host enzymatic activity (Kalaitzis & Moore, 2004). Interestingly, the phaeochromycins are produced by the wild-type strain LL-P018, illustrating that these types of unusually cyclized polyketides can occur as natural products of polyketide biosynthesis.

Structural comparison to shunt products in the biosynthesis of well-known aromatic polyketides suggested that the phaeochromycins themselves may be partially cyclized intermediates or shunt metabolites derived from a biosynthetic pathway for which the fully cyclized end product has not yet been identified. Because these compounds have importance as potential anti-inflammatory therapeutic agents, further understanding of their biosynthesis was sought. To this end, strain LL-P018 was cultured in a variety of liquid media in search of related metabolites. Also, the genome of strain LL-P018 was screened for putative phaeochromycin biosynthetic genes.

Alnumycin (Figure 4.2) (Bieber *et al.*, 1998) was identified as a metabolite of strain LL-P018. This compound was intriguing as a possible product of phaeochromycin biosynthesis due to its aromatic polyketide structure, as well as the presence of a propyl side chain which is characteristic of the phaeochromycins (Figure 4.1) (Graziani *et al.*,

2005). Presumably, the core aromatic structure of this molecule can be derived from cyclization and aromatization of an octaketide chain, which would be consistent with the proposed biosynthesis of the phaeochromycins along an actinorhodin-like pathway (Fernandez-Moreno *et al.*, 1992; Graziani *et al.*, 2005; Kalaitzis & Moore, 2004). Potential mechanisms for biosynthesis of the phaeochromycins (Appendix 11) and alnumycin (Appendix 12) are proposed in this thesis. Elucidating the actual biosynthetic pathway for these compounds will be the focus of further experiments.

Several type II PKS genes were identified from the genome of strain LL-P018 (Figure 4.3, Table 4.1). In order to determine whether these genes play a role in phaeochromycin biosynthesis, and to test the hypothesis that the phaeochromycins and alnumycin are the products of a single biosynthetic gene cluster, one of these genes was disrupted by insertion of an apramycin resistance marker (Figure 4.4). A plasmid construct was designed containing the marker inserted into the center of the KSβ gene. The marker was flanked on each side by 1kb of sequence, thus designed to allow integration of the construct into the LL-P018 genome *via* single or double-crossover homologous recombination (Figure 4.5). After transfer of the construct to LL-P018 *via* conjugation with *E. coli*, several mutants have been identified which no longer produce alnumycin or the phaeochromycins (Figure 4.6). This suggests that the type II polyketide synthase cluster identified in strain LL-P018 is responsible for the biosynthesis of all of these compounds.

Evidence suggests that the knockout construct was inserted into the native PKS *via* single-crossover homologous recombination, successfully disrupting transcription of the biosynthetic gene cluster (Figure 4.6, Appendices 8 and 9). Repeated passes of these mutants in culture through several generations may eventually result in a second crossover event, leading to double-crossover integration of the disrupted KSβ gene and elimination of the wild-type copy (Kieser *et al.*, 2000). This genotype has not yet been detected following several subcultures of knockout mutants, and efforts to achieve it are ongoing. However, the single-crossover integration successfully disrupted biosynthesis of the phaeochromycins and alnumycin, supporting our hypotheses regarding the genetic origins of these metabolites.

The set of genes identified thus far are homologous to highly conserved minimal PKS genes involved in the biosynthesis of many aromatic polyketides (Table 4.1), supporting the earlier hypothesis that the phaeochromycins are assembled along a biosynthetic route similar to that of actinorhodin (Graziani *et al.*, 2005). Interestingly, all of the genes found were closely related to genes from the frenolicin and R1128 gene clusters (Bibb *et al.*, 1994; Marti *et al.*, 2000). Like the phaeochromycins, frenolicin and R1128 are aromatic polyketides primed with non-acetate starter units (Tang *et al.*, 2003). Frenolicin, in particular, has a propyl side-chain derived from a butyrate starter unit (Hopwood, 1997), similar to that of the phaeochromycins.

Starter unit selectivity and specificity in type II polyketides has been the subject of several recent studies (Bao *et al.*, 1999; Bisang *et al.*, 1999; Tang *et al.*, 2003, 2004). In

aromatic polyketides primed with acetate, such as actinorhodin, the KSβ (also known as the chain-length factor, or CLF) catalyzes the decarboxylation of malonyl-ACP, yielding an acetate starter unit for initiation of the polyketide chain (Bisang *et al.*, 1999). The frenolicin and R1128 gene clusters contain an extra set of PKS genes which override acetate priming and allow initiation with unusual starter units (Tang *et al.*, 2003). These unusual initiation modules have shown promise as tools for biosynthetic engineering of new aromatic polyketides *via* co-expression of initiation genes from one pathway with minimal PKS genes from another (Tang *et al.*, 2004). The high degree of similarity between the phaeochromycin minimal PKS and those for frenolicin and R1128 biosynthesis, as well as structural evidence of a butyrate starter unit in phaeochromycin biosynthesis, suggest that the phaeochromycin gene cluster may also contain a unique set of initiation genes, which furthermore could have potential utility in biosynthetic engineering.

The phaeochromycins represent a new family of biologically active polyketides from *Streptomyces* (Graziani *et al.*, 2005). Cultural and genetic studies have identified a gene cluster involved in the biosynthesis of the phaeochromycins and the related polyketide alnumycin. Cloning of the entire phaeochromycin gene cluster will yield further insight into the biosynthesis of these important molecules, and may ultimately lead to the production of new phaeochromycins and alnumycins *via* biosynthetic engineering.

Acknowledgements

The authors would like to acknowledge Kerry Kulowski for molecular biology advice and assistance, Jan Kieleczawa for DNA sequencing, Nathan Magarvey for providing plasmid pJV177, and Laurie Barbieri and Leonard McDonald for chemistry advice and assistance.

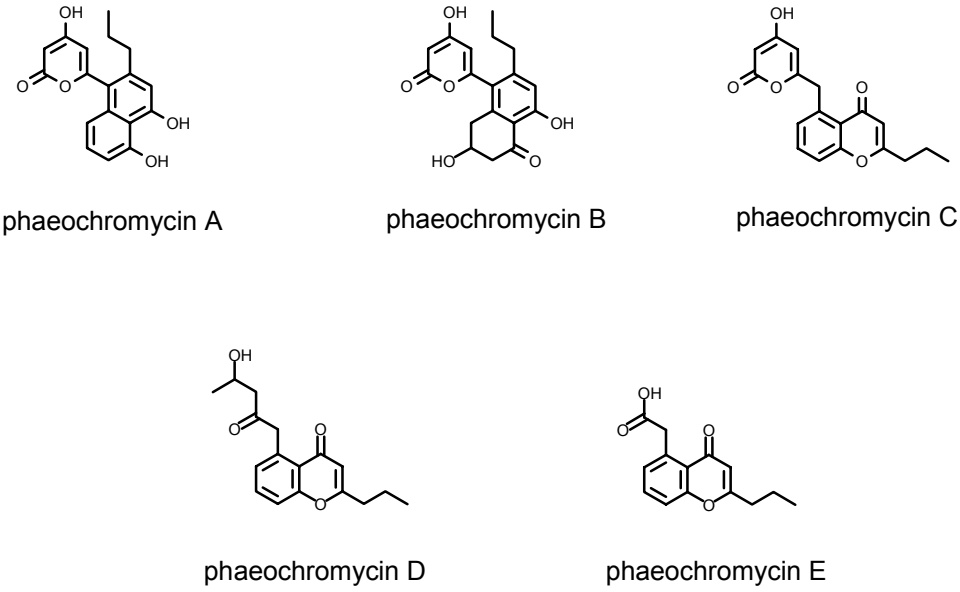


Figure 4.1. Polyketide structures of the phaeochromycins (Graziani *et al.*, 2005).

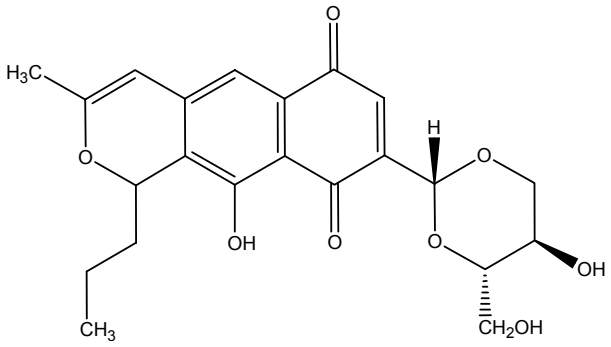


Figure 4.2. Structure of the polyketide alnumycin (Bieber *et al.*, 1998).

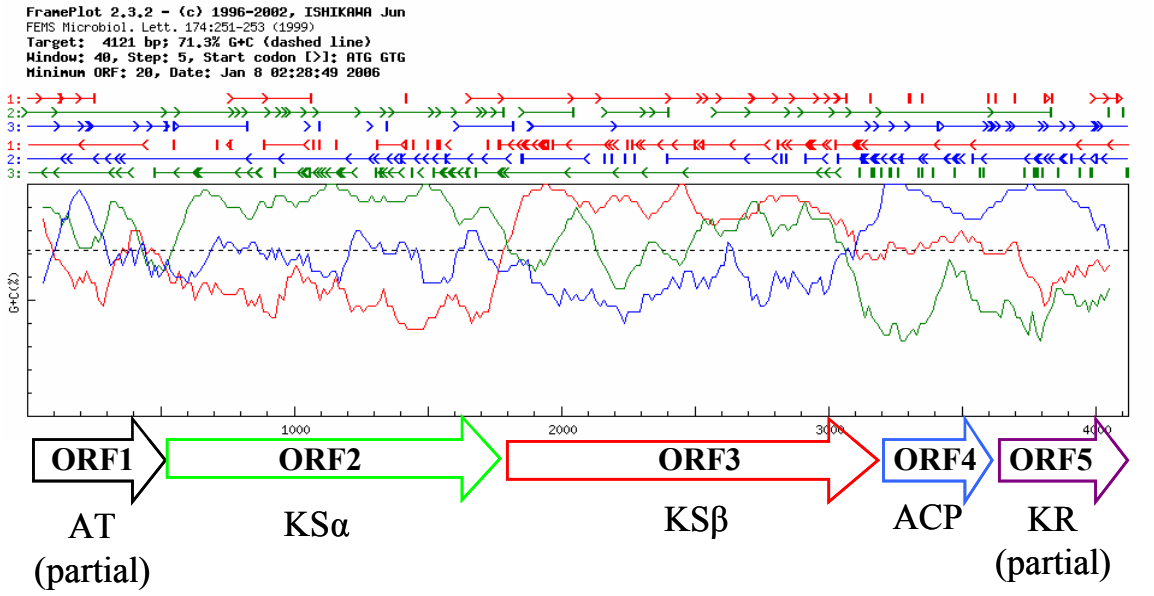


Figure 4.3. FRAME analysis and annotation of type II PKS genes from strain LL-P018. FRAME analysis and BLASTX (Altschul *et al.*, 1997) searches of open reading frames (ORFs) were performed using FramePlot (Ishikawa & Hotta, 1999). AT= acyltransferase, KSα= ketosynthase alpha subunit, KSβ= ketosynthase beta subunit (or chain length factor), ACP= acyl carrier protein, KR= ketoreductase.

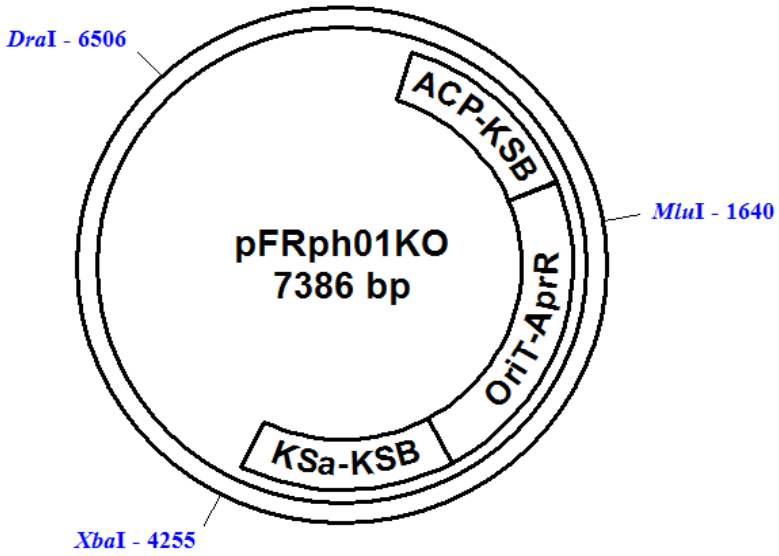


Figure 4.4. Map of plasmid pFRPH01KO. Location of restriction sites used to determine presence and orientation of insert are shown. AT= acyltransferase, KSα= ketosynthase alpha subunit, KSβ= ketosynthase beta subunit (or chain length factor), ACP= acyl carrier protein, KR= ketoreductase, OriT= origin of transfer, AprR= apramycin resistance gene. The OriT-AprR cassette is in reverse orientation relative to the PKS genes.

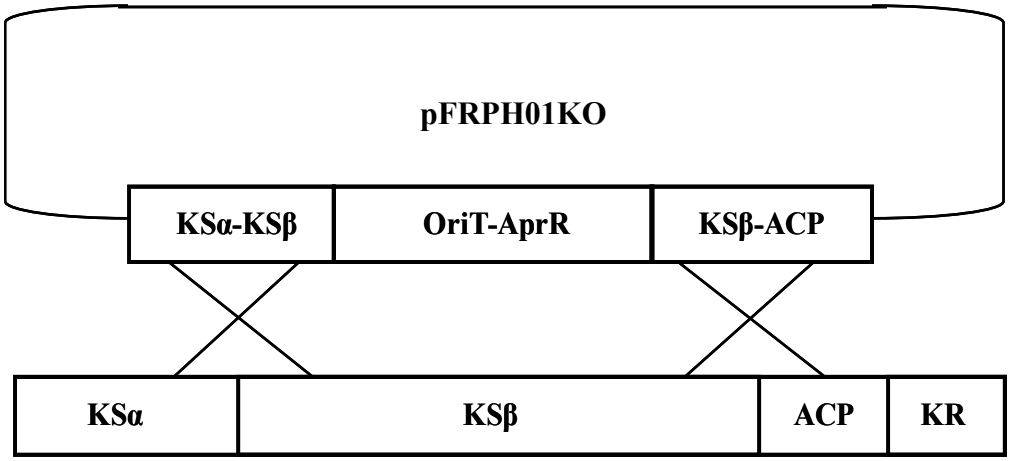


Figure 4.5. Proposed disruption of wild type KSβ gene *via* homologous recombination. Plasmid pFRPH01 contained a portion of a polyketide pathway from the chromosome of strain LL-P018. The beta-ketosynthase (KSβ) gene on plasmid pFRPH01 was disrupted by insertion an origin of transfer (OriT) and an apramycin resistance marker (AprR). The plasmid was transferred to strain LL-P018 *via* conjugation with *E. coli*, and is unable to replicate in *Streptomyces.* The construct was designed to allow insertion of the disrupted copy into the genome by single or double-crossover homologous recombination, at either or both of the homologous flanking regions shown above. Results of PCR analysis (Appendices 8 and 9) indicated that the insertion occurred *via* single-crossover. AT= acyltransferase, KSα= alpha-ketosynthase, KSβ= beta-ketosynthase (or chain length factor), ACP= acyl carrier protein, KR= ketoreductase, OriT= origin of transfer, AprR= apramycin resistance gene.

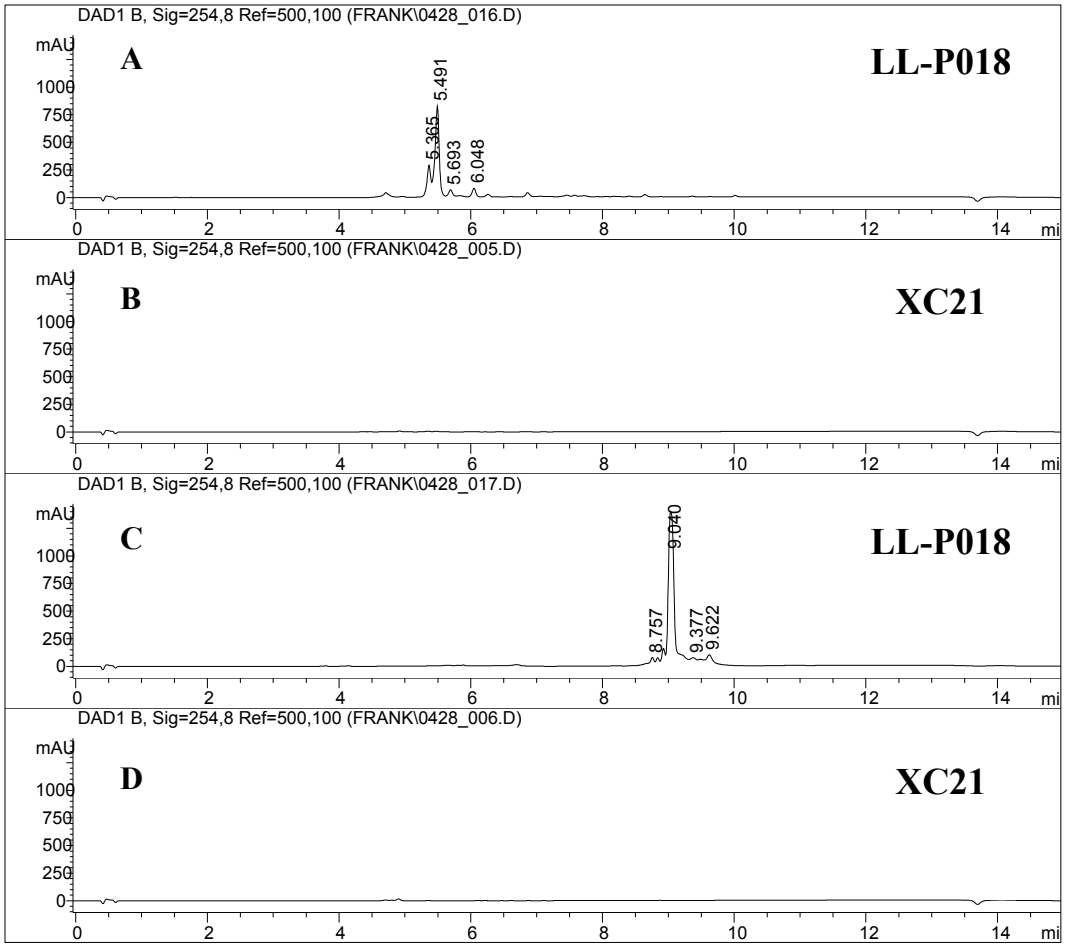


Figure 4.6. HPLC analysis of phaeochromycin and alnumycin production by strain LL-P018 and knockout mutant XC21. A: Strain LL-P018 cultured in phaeochromycin production medium. Retention times (approximate, in minutes): phaeochromycin A = 5.5, phaeochromycin B = 5.4, phaeochromycin C = 5.7. B: Knockout mutant XC21 cultured in phaeochromycin production medium. Production of phaeochromycins was not detected. C: Strain LL-P018 cultured in alnumycin production medium. Retention time (approximate, in minutes): alnumycin = 9.0. D: Knockout mutant XC21 cultured in alnumycin production medium. Production of alnumycin was not detected.

ORF	Size (bp)	Proposed function	Closest homologous protein	Gene cluster	Function
ORF1	500 (incomplete)	acyltransferase	hedF, 58% identities, 63% positives	hedamycin	putative acyl transferase
ORF2	1215	α-ketosynthase	frnL, 76% identities, 85% positives	frenolicin	ketoacyl synthase
ORF3	1287	β-ketosynthase (chain length factor)	frnM, 69% identities, 78% positives	frenolicin	chain length factor
ORF4	255	acyl carrier protein	frnN, 59% identities, 71% positives	frenolicin	acyl carrier protein
ORF5	690 (incomplete)	ketoreductase	ORF5, 69% identities, 80% positives	unknown PKS from *S. antibioticus* ATCC 11891	putative keto reductase

Table 4.1. Proposed functions of phaeochromycin open reading frames (ORFs) identified from strain LL-P018 (Figure 4.3). Open reading frames were identified, translated, and compared to protein databases using FramePlot (Ishikawa & Hotta, 1999) and BLAST (Altschul *et al.*, 1997). Protein functions were deduced based on homology to known type II PKS enzymes (Bibb *et al.*, 1994; Bililign *et al.*, 2004; Colombo *et al.*, 2001).

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Bao, W., Sheldon, P. J., Wendt-Pienkowski, E. & Hutchinson, C. R. (1999). The *Streptomyces peucetius* dpsC gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide. *J Bacteriol* **181**, 4690-4695.

Bibb, M. J., Sherman, D. H., Omura, S. & Hopwood, D. A. (1994). Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin. *Gene* **142**, 31-39.

Bieber, B., Nuske, J., Ritzau, M. & Grafe, U. (1998). Alnumycin a new naphthoquinone antibiotic produced by an endophytic *Streptomyces* sp. *J Antibiot (Tokyo)* **51**, 381-382.

Bililign, T., Hyun, C. G., Williams, J. S., Czisny, A. M. & Thorson, J. S. (2004). The hedamycin locus implicates a novel aromatic PKS priming mechanism. *Chem Biol* **11**, 959-969.

Bisang, C., Long, P. F., Cortes, J., Westcott, J., Crosby, J., Matharu, A. L., Cox, R. J., Simpson, T. J., Staunton, J. & Leadlay, P. F. (1999). A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* **401**, 502-505.

Colombo, V., Fernandez-de-Heredia, M. & Malpartida, F. (2001). A polyketide biosynthetic gene cluster from *Streptomyces antibioticus* includes a LysR-type transcriptional regulator. *Microbiology* **147**, 3083-3092.

Fernandez-Moreno, M. A., Martinez, E., Boto, L., Hopwood, D. A. & Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J Biol Chem* **267**, 19278-19290.

Flett, F., Mersinias, V. & Smith, C. P. (1997). High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* **155**, 223-229.

Graziani, E. I., Ritacco, F. V., Bernan, V. S. & Telliez, J. B. (2005). Phaeochromycins A-E, anti-inflammatory polyketides isolated from the soil actinomycete *Streptomyces phaeochromogenes* LL-P018. *J Nat Prod* **68**, 1262-1265.

Hopwood, D. A. (1997). Genetic contributions to understanding polyketide synthases. *Chem Rev* **97**, 2465-2497.

Ishikawa, J. & Hotta, K. (1999). FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol Lett* **174**, 251-253.

Kalaitzis, J. A. & Moore, B. S. (2004). Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* **67**, 1419-1422.

Keatinge-Clay, A. T., Maltby, D. A., Medzihradszky, K. F., Khosla, C. & Stroud, R. M. (2004). An antibiotic factory caught in action. *Nat Struct Mol Biol* **11**, 888-893.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*.The John Innes Foundation, Norwich, England.

Kotlyarov, A., Neininger, A., Schubert, C., Eckert, R., Birchmeier, C., Volk, H. D. & Gaestel, M. (1999). MAPKAP kinase 2 is essential for LPS-induced TNF-alpha biosynthesis. *Nat Cell Biol* **1**, 94-97.

Malpartida, F. & Hopwood, D. A. (1984). Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature (London)* **309**, 462-464.

Marti, T., Hu, Z., Pohl, N. L., Shah, A. N. & Khosla, C. (2000). Cloning, nucleotide sequence, and heterologous expression of the biosynthetic gene cluster for R1128, a non-steroidal estrogen receptor antagonist. Insights into an unusual priming mechanism. *J Biol Chem* **275**, 33443-33448.

McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A. & Khosla, C. (1994a). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci USA* **91**, 11542-11546.

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1994b). Engineered biosynthesis of novel polyketides: actVII and actIV genes encode aromatase and cyclase enzymes, respectively. *J Am Chem Soc* **116**, 10855-10859.

Metsa-Ketela, M., Salo, V., Halo, L., Hautala, A., Hakala, J., Mantsala, P. & Ylihonko, K. (1999). An efficient approach for screening minimal PKS genes from *Streptomyces*. *FEMS Microbiol Lett* **180**, 1-6.

Naruse, N., Goto, M., Watanabe, Y., Terasawa, T. & Dobashi, K. (1998). K1115 A, a new anthraquinone that inhibits the binding of activator protein-1 (AP-1) to its recognition sites. II. Taxonomy, fermentation, isolation, physico-chemical properties and structure determination. *J Antibiot (Tokyo)* **51**, 545-552.

Piel, J., Hertweck, C., Shipley, P. R., Hunt, D. M., Newman, M. S. & Moore, B. S. (2000). Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate "*Streptomyces maritimus*": evidence for the derailment of an aromatic polyketide synthase. *Chem Biol* **7**, 943-955.

Ritacco, F. V. & Eveleigh, D. E. (2007). Molecular and phenotypic comparison of phaeochromycin-producing strains of *Streptomyces phaeochromogenes* and *Streptomyces ederensis*. *Manuscript in preparation*.

Sambrook, J., Fritsch, E. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual (Second Edition)*, Second edn.Cold Spring Harbor Laboratory Press, Woodbury, NY.

Tang, Y., Lee, T. S. & Khosla, C. (2004). Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. *PLoS Biol* **2**, E31.

Tang, Y., Lee, T. S., Kobayashi, S. & Khosla, C. (2003). Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity. *Biochemistry* **42**, 6588-6595.

Xiang, L., Kalaitzis, J. A. & Moore, B. S. (2004). *EncM*, a versatile enterocin biosynthetic enzyme involved in Favorskii oxidative rearrangement, aldol condensation, and heterocycle-forming reactions. *Proc Natl Acad Sci USA* **101**, 15609-15614.

Xiang, L., Kalaitzis, J. A., Nilsen, G., Chen, L. & Moore, B. S. (2002). Mutational analysis of the enterocin favorskii biosynthetic rearrangement. *Org Lett* **4**, 957-960.

Zhang, H.-I., He, X.-G., Adefarati, A., Galucci, J., Cole, S., Beale, J., Keller, P., Chang, C.-J. & Floss, H. G. (1990). Mutactin, a Novel Polyketide from *Streptomyces coelicolor*. Structure and Biosynthetic Relationshp to Actinorhodin. *J Org Chem* **55**, 1682-1684.

Chapter 5

General discussion

The phaeochromycins (Graziani *et al.*, 2005) are a new family of bioactive polyketides from *Streptomyces*, and therefore represent potential taxonomic and biosynthetic novelty. Understanding the taxonomy of the producing microorganism and the steps in phaeochromycin biosynthesis will set these promising microbial metabolites into evolutionary and metabolic perspective, and will broaden and strengthen our knowledge of *Streptomyces*.

Streptomyces are prolific producers of complex secondary metabolites, and have been studied thoroughly for over a century (Beijerinck, 1900; Rullmann, 1895; Waksman, 1959; Waksman & Curtis, 1916). However, it is highly likely that considerable undiscovered biological and chemical novelty remains to be illuminated in *Streptomyces* (Hopwood, 2007). This is evidenced by the recent sequencing of the *Streptomyces coelicolor* genome (Bentley *et al.*, 2002), which revealed the presence of 18 new secondary metabolite gene clusters likely encoding the biosynthesis of unknown natural products, in addition to the four known pathways which had previously been identified from this organism. The *Streptomyces avermitilis* genome has also been fully sequenced and annotated, yielding similar results regarding secondary metabolic pathways (Ikeda *et al.*, 2003).

The phaeochromycins are further proof of the continued discovery of chemical novelty from the genus *Streptomyces*. With the emergence of newer and more sensitive screening technologies, certain compounds, though lacking strong pigmentation or antimicrobial activity, can be detected in microbial culture extracts based on their interaction with specific molecular targets. The phaeochromycins were not detected by traditional antimicrobial screening. The ability of these compounds to inhibit the kinase MK-2 led to their discovery in a crude culture broth extract of strain LL-P018 (Graziani *et al.*, 2005). Further analysis of related strains revealed that these novel compounds are also produced by the *Streptomyces phaeochromogenes* type strain, a species first described 90 years ago (Conn, 1917). The value of my investigation is, in part, that novel and pharmacologically important metabolites are still being discovered from some of the earliest known streptomycetes.

The taxonomy of *Streptomyces* has been difficult and confusing since its inception (Cohn, 1875; Conn, 1917; Hopwood, 2007). The patenting of antibiotic-producing streptomycetes beginning in the 1940s led to a dramatic and perhaps fortuitous increase in the number of species. These numbers have decreased over time with use of newer, more definitive standards for speciation and classification (Anderson & Wellington, 2001; Shirling & Gottlieb, 1966). Even so, the precise taxonomic placement of a streptomycete using any particular set of methods may still prove difficult. However, much can be learned about the organism and the secondary metabolites it produces by taxonomic investigation, as is detailed in Chapter 3. For instance, comparison of

metabolite production by closely related strains may lead to the discovery of new metabolites, or yield insights into biosynthetic and metabolic distinctions between strains.

The taxonomy of "new" antibiotic-producing streptomycetes is perhaps investigated with the satisfying objective of defining new species of *Streptomyces,* while also yielding potentially patentable taxonomic and chemical novelty. This did not prove true for the phaeochromycin producer strain LL-P018, as it was found to be closely related to the type strain of S. *phaeochromogenes*. However, the status of the type strain was greatly refined.

Novelty is disclosed in diverse manners. Strain variation is important amongst *Streptomyces*, as different strains of the same species can differ in various traits, including secondary metabolite production. The taxonomic study (Chapter 3) did not yield new *Streptomyces* species, yet it greatly clarified strain variation. Seven of the strains represented members of the species *S. phaeochromogenes* (Figure 3.1), but it was illustrated that there were two distinct strain groups within the species (Figures 3.4, 3.5, and 3.7). These two subgroups produce the phaeochromycin family of metabolites, but the regulation of their production appears to differ, based on their metabolite profiles in certain culture media (Figure 3.5). Furthermore, *Streptomyces ederensis* strains were not different enough from *S. phaeochromogenes* to warrant separation as a distinct species, and should be renamed as strains of *S. phaeochromogenes*.

Defining the taxonomic status of strain LL-P018 led to the development of novel approaches for strain comparison and delineation in *Streptomyces*. Cluster analysis of metabolite profiles was one useful approach (Figure 3.5). While genetic fingerprinting has proven to be a valuable means of discriminating between strains of *Streptomyces*, it does not always correlate with secondary metabolite production (Ritacco *et al.*, 2003). Metabolite profiling can effectively group strains based on their chemical fingerprint (Figure 3.5), assuming that environmental and chromatographic consistency among samples is maintained, and that data handling issues such as background subtraction and elimination of non-diagnostic peaks are carefully considered. Furthermore, the combination of metabolite profiling with genetic fingerprinting into a single, integrated analysis (Figure 3.7) represents a novel means of strain comparison based upon both genotypic and phenotypic traits. This method was effective in defining distinct strain clusters within *S. phaeochromogenes*, revealing two subgroups within the species, and illuminating outlying strains.

Study of the biosynthesis of the phaeochromycins yielded further novelty, as illustrated in Chapter 4. Based on structural characteristics, it was suspected that these novel compounds were intermediates or shunt products in the biosynthesis of a fully cyclized aromatic polyketide, which had not yet been identified. Growth of strain LL-P018 on a range of different culture media (Appendix 10a) led to the identification of alnumycin (Bieber *et al.*, 1998), a likely end product of phaeochromycin biosynthesis (Figures 4.2 and 4.6). While alnumycin is not novel, its biosynthesis and the gene cluster encoding it

have not previously been detailed. This represents a clear finding of biosynthetic novelty, as predicted in the original hypothesis.

The discovery that a single gene cluster controls biosynthesis of the phaeochromycins and alnumycin is described in Chapter 4. Transfer of knockout construct pFRPH01KO (Figure 4.4) to strain LL-P018 and integration *via* single-crossover appears to have successfully disrupted both phaeochromycin and alnumycin biosynthesis (Figure 4.6). However, the expected double-crossover homologous recombination has not yet been detected (Figure 4.5; Appendices 8 and 9). This final expected genotype would result in complete elimination of the wild-type beta-ketosynthase (KSβ) and provide the most convincing evidence of the role of this newly identified gene cluster. Efforts to identify and isolate a double-crossover mutant are ongoing.

Analysis of the phaeochromycin gene cluster will not be complete until the entire gene cluster is cloned, sequenced, and annotated. This will ultimately reveal detailed comparative data. As described in Chapter 4, type II polyketide synthases (PKSs) that are primed with non-acetate starter units, such as the phaeochromycins, typically possess an extra set of PKS genes involved in starter unit selectivity (Bisang *et al.*, 1999; Tang *et al.*, 2003, 2004). Such type II PKS pathways have been particularly useful in combinatorial biosynthesis, a type of genetic engineering in which certain genes from one biosynthetic pathway are co-expressed with certain genes from another, leading to the biosynthesis of new compounds (Hutchinson, 1999; Tang *et al.*, 2004). Sequencing the full phaeochromycin biosynthetic pathway will reveal new genes, potentially enabling

biosynthetic engineering for the production of new phaeochromycins, alnumycins, or other compounds.

While this thesis is by no means a final analysis of the phaeochromycins or strain LL-P018, it represents the diversity of the processes that are required to fully understand an organism, its taxonomic placement and its biosynthetic capabilities. Investigation into these new, biologically active polyketides, the phaeochromycins, has revealed a wealth of information, while raising a host of new taxonomic, evolutionary and biosynthetic questions.

References

Anderson, A. S. & Wellington, E. M. H. (2001). The taxonomy of *Streptomyces* and related genera. *Int J Syst Evol Microbiol* **51**, 797-814.

Beijerinck, M. W. (1900). Ueber Chinonbildung durch *Streptothrix chromogena* und Lebensweise dieses Mikroben. *Centr Bakteriol Parasitenk Abt II* **6**, 2-12. Cited in: Waksman (1959) and Waksman and Curtis (1916) .

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J. & Hopwood, D. A. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.

Bieber, B., Nuske, J., Ritzau, M. & Grafe, U. (1998). Alnumycin a new naphthoquinone antibiotic produced by an endophytic *Streptomyces* sp. *J Antibiot (Tokyo)* **51**, 381-382.

Bisang, C., Long, P. F., Cortes, J., Westcott, J., Crosby, J., Matharu, A. L., Cox, R. J., Simpson, T. J., Staunton, J. & Leadlay, P. F. (1999). A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* **401**, 502-505.

Cohn, F. (1875). Untersuchungen über Bakterien [Investigations on Bacteria] II. *Beiträge zur Biologie der Pflanzen* **1**, 141-204. Cited in: Hopwood (2007).

Conn, H. J. (1917). Soil Flora Studies. V. Actinomycetes in Soil. *NY Agr Exp Sta Tech Bull* **60**, 1-16.

Graziani, E. I., Ritacco, F. V., Bernan, V. S. & Telliez, J. B. (2005). Phaeochromycins A-E, anti-inflammatory polyketides isolated from the soil actinomycete *Streptomyces phaeochromogenes* LL-P018. *J Nat Prod* **68**, 1262-1265.

Hopwood, D. A. (2007). *Streptomyces in Nature and Medicine: The Antibiotic Makers.*Oxford University Press, Oxford, UK.

Hutchinson, C. R. (1999). Combinatorial biosynthesis of antibiotics. In *Drug Discovery from Nature*, pp. 233-254. Eds. S. Grabley & R. Thiericke. Springer-Verlag, Berlin, Germany.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. (2003). Complete genome sequence and comparative

analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**, 526-531.

Ritacco, F. V., Haltli, B., Janso, J. E., Greenstein, M. & Bernan, V. S. (2003). Dereplication of *Streptomyces* soil isolates and detection of specific biosynthetic genes using an automated ribotyping instrument. *J Ind Microbiol Biotechnol* **30**, 472-479.

Rullmann, W. (1895). Chemisch-bakteriologische Untersuchungen von Zwischendeckenfüllungen, mit besonderer Berücksichtigung von *Cladothrix odorifera.* Munich, Germany: Universität München. Cited in: Waksman and Curtis (1916) and Waksman (1959).

Shirling, E. B. & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313-340.

Tang, Y., Lee, T. S. & Khosla, C. (2004). Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. *PLoS Biol* **2**, E31.

Tang, Y., Lee, T. S., Kobayashi, S. & Khosla, C. (2003). Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity. *Biochemistry* **42**, 6588-6595.

Waksman, S. A. (1959). *The Actinomycetes, Vol.1: Nature, Occurrence, and Activities*.The Williams & Wilkins Co., Baltimore, MD.

Waksman, S. A. & Curtis, R. E. (1916). The actinomyces of the soil. *Soil Sci* **1**, 99-134.

Appendices

NRRL B16392	100																
NRRL B1131	96.9	100															
NRRL B1517	95.2	94.9	100														
NRRL B2123	94.9	94.9	99.1	100													
NRRL WC3776	95.3	96.0	97.3	97.4	100												
NRRL B2031	95.6	96.1	97.4	97.4	99.9	100											
NRRL B3559	95.4	96.4	97.2	97.3	98.8	98.7	100										
NRRL B5333	95.6	96.4	97.2	97.0	98.5	98.7	99.8	100									
NRRL B12497	94.6	95.3	95.4	96.4	96.5	96.6	95.9	96.2	100								
ISP 5170	94.4	95.1	95.7	96.2	96.5	96.8	95.9	96.2	98.6	100							
NRRL B8146	95.2	95.6	96.6	97.0	96.9	97.0	96.4	96.5	99.0	100	100						
NRRL B5478	95.6	96.1	96.9	97.5	97.1	97.2	96.9	97.0	99.0	100	100	100					
P018	94.6	95.2	95.6	96.0	96.2	96.5	96.4	96.7	98.3	99.3	99.6	100	100				
ATCC 23945	94.5	95.3	95.6	96.4	96.5	96.6	96.9	97.2	98.3	99.0	99.5	100	99.5	100			
ATCC 3338	94.5	95.3	95.6	96.4	96.5	96.6	96.9	97.2	98.3	99.0	99.5	100	99.5	100	100		
NRRL B1248	94.5	95.3	95.6	96.4	96.5	96.6	96.9	97.2	98.3	99.0	99.5	100	99.5	100	100	100	
A.utahensis	87.2	87.1	88.1	88.2	87.5	87.4	87.3	87.3	87.5	86.7	87.5	88.1	87.4	86.8	86.8	86.8	100

Appendix 1. Similarity matrix depicting 16S rRNA sequence comparison of *S. phaeochromogenes* and closely related strains. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phylogenetic tree in Chapter 3, Figure 3.2.

NRRL B1517	100															
NRRL B3559	83.4	100														
NRRL B5333	77.0	75.9	100													
NRRL B1131	71.7	70.2	52.8	100												
ISP 5170	67.5	43.9	38.9	60.6	100											
NRRL B2123	79.4	66.1	67.2	60.7	71.6	100										
NRRL B16392	52.1	57.7	47.2	48.2	50.3	56.0	100									
NRRL B2031	38.7	39.9	41.1	31.1	64.2	54.2	65.4	100								
ATCC 3338	15.8	50.0	27.6	22.2	32.0	21.7	28.3	49.9	100							
NRRL B1248	18.9	51.8	30.7	29.1	29.7	24.3	27.3	50.2	98.5	100						
ATCC 23945	10.9	49.0	24.3	13.5	21.7	16.3	32.8	44.8	97.9	95.3	100					
P018	27.7	56.4	44.4	29.5	34.1	19.8	30.0	57.0	92.6	91.6	91.8	100				
NRRL B5478	41.8	51.3	53.1	23.4	33.1	26.5	49.2	46.2	76.7	72.0	81.9	82.1	100			
NRRL B8146	27.6	47.2	42.5	16.9	23.7	19.0	45.8	39.4	82.8	76.8	88.7	83.1	97.4	100		
NRRL B12497	32.3	50.7	42.7	14.9	33.8	24.2	55.5	67.5	78.0	78.3	78.7	82.5	82.0	78.8	100	
NRRL WC3776	20.1	13.6	13.6	19.7	12.1	39.5	36.3	13.9	4.7	4.3	7.8	19.9	14.9	13.8	11.4	100

Appendix 2. Similarity matrix depicting ribotype comparison of *S. phaeochromogenes* and closely related strains. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.4.

NRRL B12497	100															
NRRL B16392	81.9	100														
NRRL B3559	69.0	69.7	100													
NRRL B1517	57.2	54.2	45.0	100												
NRRL B2123	66.1	51.6	63.2	74.4	100											
NRRL B5333	43.1	40.0	67.9	50.9	48.3	100										
NRRL WC3776	47.9	27.3	40.2	33.5	48.1	40.2	100									
NRRL B5478	18.7	20.0	18.8	41.7	21.7	38.5	16.5	100								
NRRL B8146	23.1	22.0	21.6	44.2	24.5	40.8	20.2	99.2	100							
NRRL B1248	19.8	20.3	21.3	44.7	22.9	45.0	18.9	98.8	98.8	100						
ATCC 3338	19.3	21.7	20.8	40.0	20.8	40.7	16.1	97.4	97.5	97.4	100					
P018	16.1	19.2	17.8	37.2	18.8	36.8	13.4	97.4	96.7	96.3	99.6	100				
ISP 5170	18.3	20.9	19.0	40.5	21.2	38.2	15.2	99.1	97.9	97.8	98.9	99.2	100			
ATCC 23945	40.7	30.6	36.6	48.3	38.0	49.1	34.0	87.4	91.5	88.1	88.8	86.5	86.3	100		
NRRL B2031	29.6	50.0	28.7	10.0	6.1	5.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	
NRRL B1131	2.5	0.0	4.8	0.0	6.9	11.7	33.3	0.5	2.0	3.2	1.9	0.4	0.0	9.1	0.0	10(

Appendix 3. Similarity matrix depicting metabolite profile comparison of *S. phaeochromogenes* and closely related strains cultured in liquid medium 1. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.5a.

NRRL B1517	100															
NRRL B5333	46.9	100														
ATCC 3338	25.9	35.5	100													
NRRL B1248	26.4	36.8	99.0	100												
ATCC 23945	24.8	36.9	98.7	98.0	100											
P018	19.9	31.0	96.3	95.5	96.9	100										
NRRL B1131	16.2	15.7	3.8	3.3	2.9	1.6	100									
NRRL B2123	45.3	15.5	13.9	14.1	13.0	12.9	41.7	100								
NRRL B12497	6.1	20.0	6.0	6.4	10.6	9.9	3.1	5.5	100							
ISP 5170	0.0	0.0	0.0	0.0	0.0	0.3	9.5	0.0	6.8	100						
NRRL B8146	0.0	0.0	4.1	4.7	2.5	18.1	3.3	0.0	1.1	86.7	100					
NRRL B5478	0.6	0.0	30.3	31.4	26.6	45.7	0.0	10.6	0.1	60.0	77.9	100				
NRRL WC3776	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.6	5.9	1.3	100			
NRRL B16392	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100		
NRRL B2031	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	64.6	100	
NRRL B3559	17.9	18.0	13.3	15.2	13.9	3.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	45.5	31.8	100

Appendix 4. Similarity matrix depicting metabolite profile comparison of *S. phaeochromogenes* and closely related strains cultured in liquid medium 2. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.5a.

NRRL B3559	100															
NRRL B5333	71.6	100														
NRRL B1131	48.2	35.6	100													
NRRL B12497	47.7	32.7	29.2	100												
NRRL WC3776	41.5	19.6	16.8	31.1	100											
NRRL B2031	30.5	0.0	0.0	0.0	20.8	100										
ATCC 3338	0.0	0.0	7.3	0.0	0.0	0.0	100									
NRRL B1248	0.0	0.0	4.3	0.0	0.0	0.0	96.0	100								
NRRL B8146	0.0	0.0	12.8	0.1	0.0	0.0	89.4	94.5	100							
ATCC 23945	0.0	0.0	0.9	0.0	0.0	0.0	90.5	90.9	83.0	100						
ISP 5170	0.0	0.0	3.0	0.0	0.0	0.0	81.8	91.6	92.1	74.7	100					
NRRL B5478	0.0	0.0	6.2	1.5	0.0	0.0	76.9	87.5	87.8	67.2	98.7	100				
P018	0.0	0.0	2.6	0.0	0.0	0.0	74.9	82.9	83.5	66.3	92.7	92.4	100			
NRRL B1517	0.0	0.0	0.0	0.0	0.0	0.0	22.2	22.0	8.5	36.8	1.3	0.0	0.0	100		
NRRL B2123	0.0	0.0	0.0	2.6	0.0	0.0	25.4	25.6	13.8	41.2	6.3	1.1	3.9	89.2	100	
NRRL B16392	0.0	0.0	0.0	0.0	0.0	0.0	13.5	20.2	13.6	11.5	13.9	19.1	8.6	0.0	0.0	100

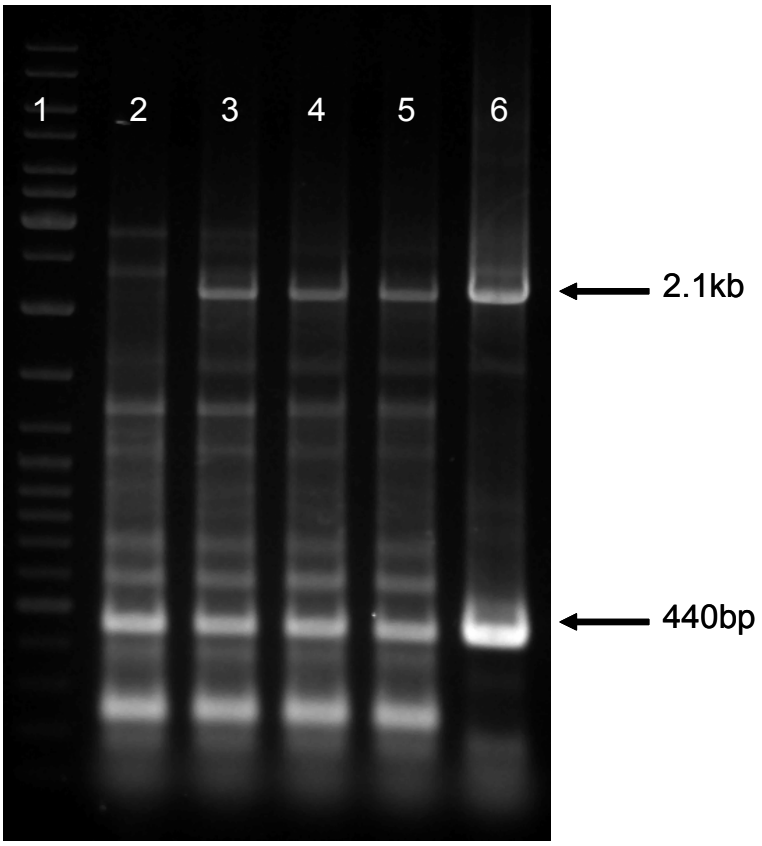
Appendix 5. Similarity matrix depicting metabolite profile comparison of *S. phaeochromogenes* and closely related strains cultured in liquid medium 3. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.5a.

NRRL B3559	100															
NRRL B5333	52.5	100														
NRRL B12497	38.9	31.9	100													
NRRL WC3776	27.2	19.9	26.3	100												
NRRL B1517	21.0	32.6	21.1	11.2	100											
NRRL B2123	21.1	21.2	24.7	16.0	69.6	100										
NRRL B16392	38.4	13.3	27.3	9.1	18.1	17.2	100									
NRRL B2031	30.3	1.9	9.9	6.9	3.3	2.0	38.2	100								
ISP 5170	6.3	12.7	8.4	13.6	14.0	9.2	11.6	0.0	100							
NRRL B8146	7.2	13.6	8.1	8.7	17.6	12.8	11.9	0.0	92.2	100						
NRRL B5478	6.3	12.8	6.8	5.9	14.1	11.1	13.0	0.0	85.9	88.3	100					
ATCC 3338	11.4	25.4	8.4	5.4	29.4	20.0	11.7	0.0	60.2	63.6	68.2	100				
NRRL B1248	12.2	27.3	8.7	6.3	31.0	20.8	13.5	0.0	63.1	66.0	72.6	97.4	100			
ATCC 23945	16.8	28.7	17.1	11.3	36.7	30.7	14.0	0.0	53.7	59.0	60.4	92.7	92.3	100		
P018	6.9	22.6	8.6	4.5	19.0	11.9	9.3	0.0	64.1	66.1	78.5	90.3	91.6	83.2	100	
NRRL B1131	17.7	21.0	11.6	16.7	5.4	16.2	0.0	0.0	4.2	6.1	2.2	4.3	3.6	4.3	1.5	100

Appendix 6. Similarity matrix depicting combined metabolite profile comparison of *S. phaeochromogenes* and closely related strains cultured in liquid media 1, 2, and 3. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.5b.

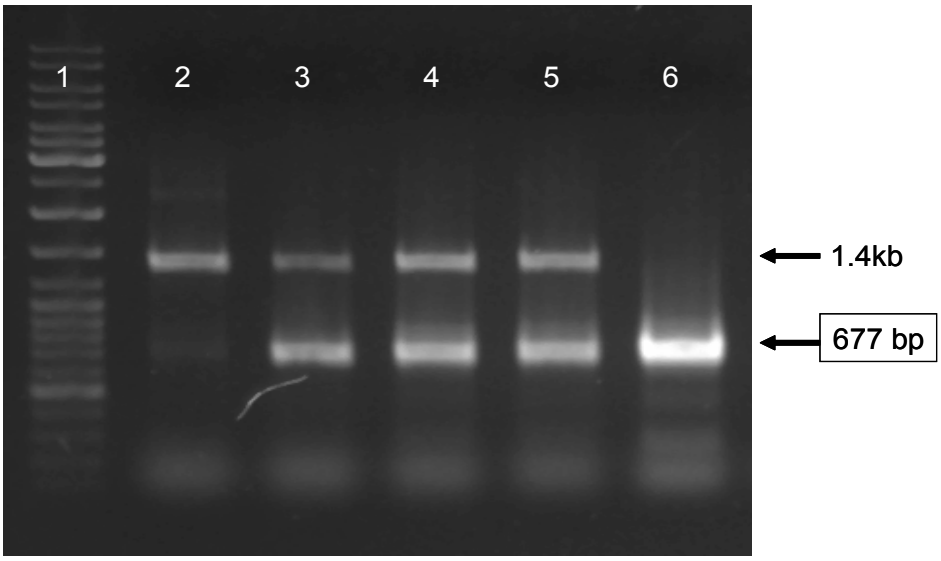
NRRL B1131	100															
NRRL WC3776	17.5	100														
NRRL B3559	30.8	23.8	100													
NRRL B5333	29.0	18.3	58.4	100												
NRRL B12497	12.4	22.6	41.9	34.6	100											
NRRL B1517	22.0	13.4	36.6	43.7	23.9	100										
NRRL B2123	27.3	21.9	32.3	32.7	24.6	72.1	100									
NRRL B16392	12.0	15.9	43.2	21.8	34.3	26.6	26.9	100								
NRRL B2031	7.8	8.7	32.7	11.7	24.3	12.2	15.1	45.0	100							
NRRL B5478	7.5	8.2	17.5	22.9	25.6	21.0	15.0	22.1	11.6	100						
NRRL B8146	8.8	10.0	17.2	20.8	25.8	20.1	14.3	20.4	9.8	90.6	100					
ISP 5170	18.3	13.2	15.7	19.3	14.7	27.4	24.8	21.3	16.1	72.7	75.1	100				
ATCC 3338	8.8	5.2	21.0	26.0	25.8	26.0	20.5	15.9	12.5	70.3	68.4	53.2	100			
NRRL B1248	10.0	5.8	22.1	28.1	26.1	28.0	21.7	17.0	12.6	72.4	68.7	54.8	97.7	100		
ATCC 23945	6.6	10.5	24.9	27.6	32.5	30.2	27.1	18.7	11.2	65.8	66.4	45.7	94.0	93.1	100	
P018	8.5	8.3	19.3	28.0	27.1	21.2	13.9	14.5	14.3	79.4	70.3	56.6	90.9	91.6	85.4	100

Appendix 7. Similarity matrix depicting combined comparison of *S. phaeochromogenes* and closely related strains by ribotype and metabolite profiles from three liquid media. Numbers represent percent similarity values. Order of columns (left to right) corresponds to order of rows (top to bottom). This comparison is depicted as a phenogram in Chapter 3, Figure 3.7.



Appendix 8. PCR screening of exconjugants for the presence of a β-ketosynthase (KSβ) disruption (Chapter 4). Lane 1= 5µL GeneRuler DNA Ladder (Fermentas). Lane 2 = wild-type strain LL-P018, 15µL of PCR mixture. Lane 3 = exconjugant strain XC20, 15µL of PCR mixture. Lane 4 = exconjugant strain XC21, 15µL of PCR mixture. Lane 5 = exconjugant strain XC22, 15µL of PCR mixture. Lane 6 = plasmid mix control, containing a 1:1 mixture of plasmids pFRPH01 and pFRPH01KO as template, 15µL of PCR mixture. Amplification of a portion of the intact KSβ gene yielded a 440bp product, while amplification of the disrupted gene yielded a 2.1kb product. The intact gene was detected in the genomic DNA of wild-type strain LL-P018, while the disrupted

gene remained undetected. Amplification from the genomic DNA of exconjugants XC20, XC21, and XC22 yielded a 400-bp and a 2.1kb PCR product, suggesting the presence of a disrupted KSβ gene along with an intact copy.



Appendix 9. PCR screening of exconjugants to determine the presence of a single-crossover or double-crossover β-ketosynthase (KSβ) disruption (Chapter 4). Lane 1= 5µL GeneRuler DNA Ladder (Fermentas). Lane 2 = wild-type strain LL-P018, 15µL of PCR reaction mixture. Lane 3 = exconjugant strain XC20, 15µL of PCR reaction mixture. Lane 4 = exconjugant strain XC21, 15µL of PCR reaction loaded. Lane 5 = exconjugant strain XC22, 15µL of PCR reaction loaded. Lane 6 = plasmid pFRPH01KO control, 15µL of PCR reaction loaded. Amplification of pCR-BluntII-TOPO (Invitrogen) backbone DNA yielded a predicted 677bp product. This product was not detected in the genomic DNA of wild-type strain LL-P018, but was detected in all three exconjugants, indicating the presence of a single-crossover disruption. An unidentified PCR product approximately 1.4kb in size was also amplified from wild-type strain LL-P018 genomic as well as all three exconjugants, but not from plasmid control DNA, and is therefore not derived from pCR-BluntII-TOPO.

Appendix 10. List of media used in this study.

Appendix 10a. Cultivation and production media

1. Agar medium A-1

For routine cultivation of *Streptomyces* strains

Ingredient	g/L
D-Glucose (dextrose)	10
Soluble starch	20
Yeast extract	5
NZ-Amine A	5
Calcium carbonate	1
Agar	15

pH adjust to 7.3

2. Liquid seed medium A-1

For *Streptomyces* liquid seed cultures

Ingredient	g/L
D-Glucose (dextrose)	10
Soluble starch	20
Yeast extract	5
NZ-Amine A	5
Calcium carbonate	1

pH adjust to 7.3

3. Liquid medium 1 (phaeochromycin production medium)

Ingredient	g/L
Magnesium sulfate ·7 H_2O	0.5
Potassium chloride	0.5
Potassium phosphate dibasic	2.5
Sodium chloride	5
Agar	0.4
Glycerol	10
Soy peptone	5

pH adjust to 7.0

4. Liquid medium 2

Ingredient	g/L
Magnesium sulfate · 7 H_2O	0.1
Potassium phosphate monobasic	0.5
D-Mannitol	10
Peptone	1
Soluble starch	15
Calcium chloride	1
MOPS pH 7.0 (2.5 M solution)	20mL

pH adjust to 7.0

5. Liquid Medium 3 (alnumycin production medium)

Ingredient	g/L
Magnesium sulfate · 7 H_2O	0.1
Potassium phosphate monobasic	0.5
D-Glucose (dextrose)	1
D-Mannitol	10
Peptone	1
Soluble starch	15
Calcium chloride	1
MOPS pH 7.0 (2.5 M solution)	20mL

pH adjust to 7.0

Appendix 10b. Taxonomic media- trace salts and carbohydrate utilization media.

6. Shirling and Gottlieb trace salts (Shirling & Gottlieb, 1966).

Ingredient	g/ 100mL
Ferrous sulfate · 7 H_2O	0.1
Manganese chloride · 4 H_2O	0.1
Zinc sulfate · 7 H_2O	0.1
Distilled water	100 mL

7. Pridham and Gottlieb ISP trace salts (Shirling & Gottlieb, 1966).

Ingredient	g/ 100mL
Cupric sulfate · 5 H_2O	0.64
Ferrous sulfate · 7 H_2O	0.11
Manganese chloride · 4 H_2O	0.79
Zinc sulfate · 7 H_2O	0.15
Distilled water	100 mL

8. Pridham and Gottlieb basal mineral salts agar (Shirling & Gottlieb, 1966).

Ingredient	g/L
Ammonium sulfate	2.64
Potassium phosphate monobasic	2.38
Potassium phosphate dibasic	5.65
Magnesium sulfate · 7H_2O	1
Agar	15
Shirling and Gottlieb trace salts	1 mL/L

pH adjust to 7.0

9. Carbohydrate stock solutions (Shirling & Gottlieb, 1966).

For taxonomic evaluation

Ingredient	g/L
D-Glucose (dextrose)	100
L-Arabinose	100
Sucrose	100
D-Xylose	100
I-Inositol	100
D-Mannitol	100
D-Fructose	100
Rhamnose	100
Raffinose	100
Cellulose (microcrystalline)	100

For taxonomic evaluation of carbohydrate utilization, add 10mL of stock solution (sterilized) to 100mL of Pridham and Gottlieb basal mineral salts agar

Appendix 10c. Taxonomic media - sporulation agars.

10. ISP2 (Shirling & Gottlieb, 1966).

Yeast extract -malt extract agar

Ingredient	g/L
Bacto yeast extract (Difco)	4
Bacto malt extract (Difco)	10
Bacto dextrose (Difco)	4
Bacto agar (Difco)	20

pH adjust to 7.3

11. ISP3 (Shirling & Gottlieb, 1966).

Oatmeal agar

Ingredient	g/L
Oatmeal	20
Agar	18

Cook or steam 20g oatmeal in 1L H20 for 20 min,
filter through cheese cloth.
Bring volume back to 1L with distilled H2O
Add 1.0 mL Pridham and Gottlieb ISP trace salts

pH adjust to 7.2

12. ISP4 (Shirling & Gottlieb, 1966).

Inorganic salts - starch agar

Ingredient	g/L
Potassium phosphate dibasic	1
Magnesium sulfate · $7H_2O$	1
Sodium chloride	1
Ammonium sulfate	2
Calcium carbonate	2
Shirling and Gottlieb trace salts	1mL/L
Agar	20

pH adjust to 7.0

13. ISP5 (Shirling & Gottlieb, 1966).

Glycerol-asparagine agar

Ingredient	g/L
L-Asparagine	1
Glycerol	10
Potassium phosphate dibasic	1
Shirling and Gottlieb trace salts	1mL/L
Agar	20

pH adjust to 7.0

Appendix 10d. Molecular biology media.

14. LB (Luria-Bertani) broth.

Ingredient	g/L
Bacto tryptone (Difco)	10
Bacto yeast extract	5
Sodium Chloride	10

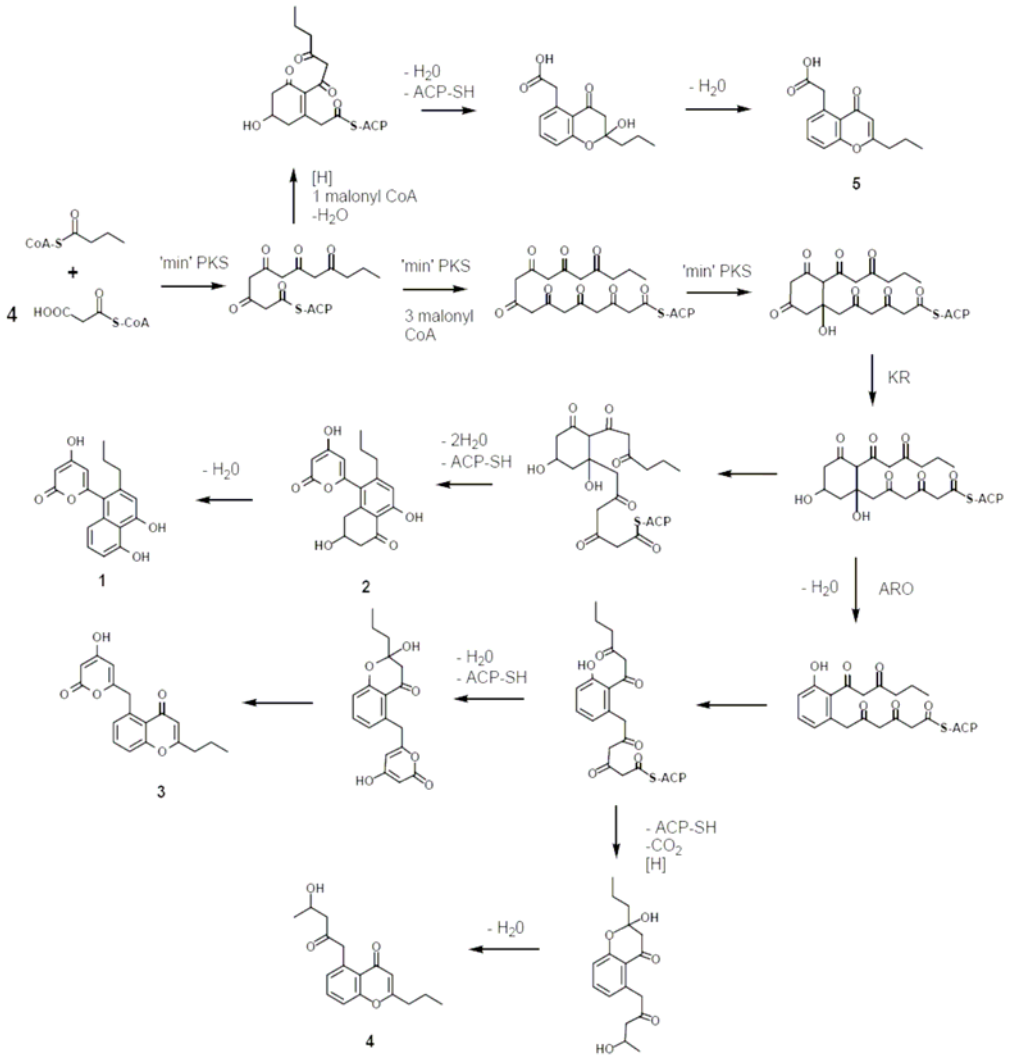
pH adjust to 7.5

15. MS (Mannitol-soya flour) agar medium (Kieser *et al.*, 2000).

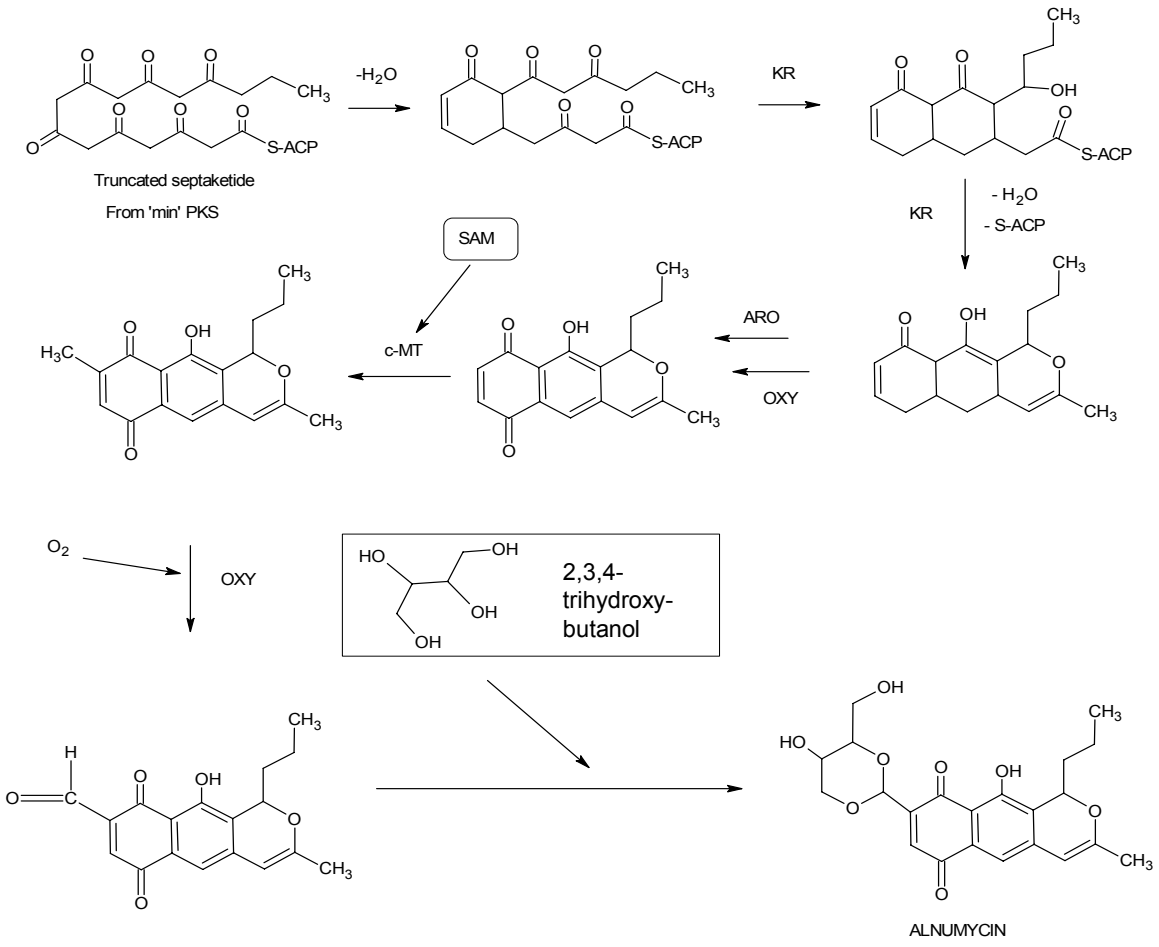
For intergeneric conjugation

Ingredient	g/L
Agar	20
Mannitol	20
Soya flour	20
Tap water	to 1L

No pH adjustment. Autoclave twice.



Appendix 11. Proposed biosynthesis of phaeochromycins A-E (**1** to **5**). Provided by Edmund Graziani (2005, personal communication). 'min' PKS = minimal polyketide synthase; KR = ketoreductase; ARO = aromatase; ACP = acyl carrier protein; CoA = coenzyme A.



Appendix 12. Proposed biosynthesis of alnumycin from the phaeochromycin polyketide chain. 'min' PKS = minimal polyketide synthase; SAM = s-adenosyl methionine; c-MT = C-methyltransferase; KR = ketoreductase; ARO = aromatase; ACP = acyl carrier protein; P_{450} = cytochrome P_{450}, OXY = oxygenase.

Culture number	Equivalent strains / Cross-reference	Source of Origin	Date accessioned	Products	Comments
LL-P018		Wyeth Research	12/18/1991	phaeochromycins A-E, alnumycin	
ATCC 3338^{T}	ATCC 23945, NRRL B-1248, CBS 288.60, CBS 929.68, DSM 40073, IFO 12898, IMET 40355, ISP 5073, JCM 4070, JCM 4659, NCIB 8505, RIA 1119	Waksman, IMRU	1963	tyrosinase, phaeochromycins A-E, alnumycin	type strain
ATCC 23945^{T}	ATCC 3338, NRRL B-1248, CBS 288.60, CBS 929.68, DSM 40073, IFO 12898, IMET 40355, ISP 5073, JCM 4070, JCM 4659, NCIB 8505, RIA 1119	Shirling from Waksman, IMRU.	1970	tyrosinase, phaeochromycins A-E, alnumycin	type strain
NRRL B-1131	RGB A-803a	R.G.Bennett, NRRL	9/1/1954	glucose isomerase	
NRRL B-1248^{T}	ATCC 23945, ATCC 3338, CBS 288.60, CBS 929.68, DSM 40073, IFO 12898, IMET 40355, ISP 5073, JCM 4070, JCM 4659, NCIB 8505, RIA 1119	ATCC	5/22/1951	glucose isomerase, phaeochromycins A-E, alnumycin	type strain
NRRL B-1517	NIHJ 108-A-1	NIHJ	12/8/1953	moldin	
NRRL B-2031	B-870	Kuroya, Tohoku U., Sendai, Japan	1/1/1958	thiomycin	
NRRL B-2123	788-A2	Okami, NIHJ, Japan	8/15/1958	probably ractinomycin A and B	
NRRL B-3559	Nakazawa C-3P, S-1544	Nakazawa, IFO	6/6/1957	chloramphenicol, glucose isomerase, restriction endonuclease Sph I	
NRRL B-5333	Strandberg NMN	Strandberg, derived from NRRL B-3559	10/16/1971	glucose-isomerase	
NRRL B-5478	ATCC 21346, FERM 229, Kai C 715-7	ATCC	5/9/1972	macarbomycin, phaeochromycins A-E, alnumycin	U.S. Pat. 3,564,090
NRRL B-16392	TUB B-233	Szakas, Tech. Univ. Budapest	2/2/1989		
ISP-5170	INA 1703/53	Shirling, ISP from Preobrazhenskaya	6/1/1964	phaeochromycins A-E, alnumycin	
WC-3776	O-263, IMRU 3776	Waksman from Okami, NIHJ	2/9/1983	chloramphenicol	
NRRL B-8146^{T}	ATCC 15305, CBS 545.70, DSM 40741, FH 2861, JCM 4958	Kutzner, DSM from Hoechst	11/11/1975	meonomycins A, B1, B2, C, phaeochromycins A-E, alnumycin	type strain, U.S. Pat. 3,674,866
NRRL B-12497^{T}	ATCC 27470, CBS 757.72, DSM 40560, IFO 13456, IMET 43541, INA 8173, ISP 5560, JCM 4837, RIA 1417	Preobrazhenskaya as INA 8173	6/24/1986	tauromycetin	type strain

Appendix 13. Provenance of strains used in this study. Presumptive culture ID and source information for these strains is listed in table 3.1.

References

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*.The John Innes Foundation, Norwich, England.

Shirling, E. B. & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313-340.

Curriculum Vita

Frank V. Ritacco

Education:

Doctoral degree candidate (2004-present), **Microbiology and Molecular Genetics**, Rutgers University / University of Medicine and Dentistry of New Jersey (joint degree program) New Brunswick, NJ.
Ph.D. Thesis: "The phaeochromycins from *Streptomyces* strain P018: from taxonomy to novelties of biosynthesis."

Master of Science Degree, **Microbiology and Molecular Genetics**,Rutgers University / University of Medicine and Dentistry of New Jersey (joint degree program) New Brunswick, NJ. October 1996.
Master's Thesis: "Development of a Chemiluminescent Gene Probe Assay for gene *hbl*A from *Bacillus cereus*."

Bachelor of Science Degree, **Biological Sciences**,
Cook College, Rutgers University, New Brunswick, NJ. May 1992.

Research Experience:

Research Scientist II: Microbiologist (July 2006-Present)
Research Scientist I: Microbiologist (January 2002-June 2006)
Scientist II: *Research Microbiologist* (June 1999-December 2001)
Scientist I: *Research Microbiologist* (October 1998-June 1999)
Discovery Research, Natural Products Microbiology Department, Wyeth Research, Pearl River, NY.

Scientist I: Research Microbiologist (May 1996-October 1998)
Bioprocess Development Department, Wyeth-Ayerst Research, Pearl River, NY.

Graduate Assistant (June 1993-May 1996)
Department of Biochemistry and Microbiology, Cook College, Rutgers University, New Brunswick, NJ.

Graduate Assistant/Lab Rotation (September 1992-May 1993)
Waksman Institute of Microbiology, Rutgers University, Piscataway, NJ.

Undergraduate Research Assistant (September 1991-May 1992)
Department of Biochemistry and Microbiology, Cook College, Rutgers University, New Brunswick, NJ.

Other Employment:

Teaching Assistant (September 1993-December 1995)
Rutgers University, New Brunswick, NJ.

Science Resident (September 1995-May 1996)
Science Residents Program, Center for Math, Science, and Computer Education
Rutgers University, Piscataway NJ.

Publications:

Graziani, E. I.; **Ritacco, F. V.;** Bernan, V. S.; Telliez, J-B. Phaeochromycins A-E, Anti-inflammatory Polyketides Isolated from the Soil Actinomycete *Streptomyces phaeochromogenes* LL-P018. Journal of Natural Products **68(8):** 1262-1265, August 2005.

Frank V. Ritacco, E. Graziani, M. Summers, T. M. Zabriskie, K. Yu, V. S. Bernan, G. Carter and M. Greenstein. Production of Novel Rapamycin Analogs by Precursor-Directed Biosynthesis. Applied and Environmental Microbiology **71(4):** 1971-1976, April 2005.

Ritacco, **F. V.,** B. Haltli, J. E. Janso, M. Greenstein and V. S. Bernan. Dereplication of *Streptomyces* soil isolates and detection of specific biosynthetic genes using an automated ribotyping instrument. Journal of Industrial Microbiology and Biotechnology **30(8):** 472-479, August 2003.

Graziani, E. I., **F. V. Ritacco**, M.Y. Summers, T. M. Zabriskie, K. Yu, V. S. Bernan, M. Greenstein, and G. T. Carter. Novel Sulfur-Containing Rapamycin Analogs Prepared by Precursor-Directed Biosynthesis. Organic Letters **5(14):** 2385-2388, 2003.

Ruzin, A., A. Severin, **F. Ritacco**, K. Tabei, G. Singh, P.A. Bradford, M.M. Siegel, S.J. Projan, and D.M. Shlaes. Further Evidence that a Cell Wall Precursor [C_{55}-MurNAc-(Peptide)-GlcNAc] Serves as an Acceptor in a Sorting Reaction. Journal of Bacteriology **184(8):** 2141-2147, April 2002.

Williams, D.E., V.S. Bernan, **F.V. Ritacco**, W.M. Maiese, M. Greenstein, and R.J. Andersen. Holyrines A and B, possible intermediates in Staurosporine biosynthesis produced in culture by a marine actinomycete derived from the North Atlantic Ocean. Tetrahedron Letters **40:** 7171-7174, 1999.