Actinomycetes sourced from unique environments as a promising source of new TB-active natural products

Miaomiao Liu

School of Natural Sciences

Griffith University

Submitted in fulfillment of the requirements of the degree of Doctor of Philosophy

Supervisors

Prof. Ronald J Quinn

Prof. Lixin zhang

Dr. Tanja Grkovic

Prof. Xueting Liu

January 2017

Abstract

Tuberculosis (TB) is the leading cause of death from infectious diseases in the world, affecting more than ten million patients each year. However, multi-drug resistance (MDR-TB) threatens progress achieved in TB care and control, and there are few drugs available to treat MDR-TB. Our overall aim was to identify anti-TB natural products from microbes sourced from unique environments. This thesis presents efforts to achieve an effective approach to identify anti-TB microbial natural products with the combination of one strain many compounds (OSMAC) strategy, NMR fingerprint and principal component analysis.

The thesis begins with an introduction of TB and the current anti-TB drugs and candidates. It also covers a review on anti-TB natural products from marine microbe and endophyte origin and analysis of their physicochemical properties using Lipinski's rule of five as well as the ChemGPS tool. As part of a research program aiming to identify anti-TB microbial constituents, a cell-based screening assay was developed to screen 2562 crude extracts. Among the active hits, 46 actinomyces isolated from marine, desert or Traditional Chinese Medicines were selected for further chemical investigation according to their chemical profiles or anti-TB activities. The results are presented in chapters 2 to 7.

Chapter 2 presents the evaluation of the potential of endophytes and marine microbial-derived natural products as anti-TB leads using the comparison of their physico-chemical properties with TB drugs and candidates. A model to predict anti-TB activity of untested compounds is also proposed.

Chapter 3 describes a cell-based HTS platform for discovering small molecules as well as crude extracts showing growth inhibition to *Mycobacteria bovis* BCG. With the aim to identify potent anti-TB natural products from actinomycetes, a microbial extract library was generated from 654 strains and 2562 crude extracts were evaluated for their activity. 415 actives were identified as the Hit Library that was the basis for this TB-focus project.

Chapter 4 demonstrates the chemical and biological investigation of an endophyte Y3111. Chemical investigation of the biota sample resulted in the identification of 6 pluramycin-type compounds, of which four were new compounds. The structures of the isolated compounds were determined by the analysis of NMR and MS data. Anti-BCG screening

suggested that 3 compounds had moderate effects on the inhibition of the bacteria.

Chapter 5 describes the isolation of 24 compounds from 21 active fractions, representing 3 endophytes, 1 desert strain and 3 marine strains. One new compound was identified from marine strain MS110154. The structures of the isolated compounds were elucidated on the basis of NMR and MS data. Four compounds exhibited significant anti-BCG activity with MIC values of 1.5, 1.5, 6 and 8 μM, respectively.

Chapter 6 presents the OSMAC investigation on 13 strains from chapter 5. NMR fingerprinting and PCA of the data set generated from 4160 fractions revealed 37 outliers, 3 of which were selected for scale-up investigations.

Chapter 7 describes the isolation and identification of 20 natural products from the 3 selected Hit fractions. One new compound with a novel structure was identified from the desert actinomycete LS120194 under the effect of a specific medium. The structures of the isolated compounds were deduced on the basis of NMR and MS data. Anti-BCG screening resulted in 8 active compounds.

The thesis concluded with an analysis of the physico-chemical properties and ChemGPS analysis of the 50 isolated microbial natural products. Most of the compounds identified in chapter 4 to 7 had druggable properties and clustered in a chemical space with current drugs. The highlighted compounds can be potentially used as lead compounds in future anti-TB drug discovery.

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

(Signed) __

      Miaomiao Liu

Table of Contents

List of Figures

List of Tables

List of Abbreviations

$[\alpha]_D$	optical rotation
1D	one dimension
2D	two dimension
3D	three dimension
BCG	Bacillus Calmette–Guérin
br	broad
BS	*Bacillus subtilis*
C_{18}	octadecyl bonded silica
$CDCl_3$	deuterated chloroform
COSY	gradient correlation spectroscopy (^{1}H-^{1}H)
d	doublet
DCM	dichloromethane
DMSO	dimethyl sulfoxide
DMSO-d_6	deuterated DMSO
DNA	deoxyribonucleic acid
DNP	dictionary of natural products
e.g.	for example
GFP	green fluorescent protein
g	gram (s)
h	hour (s)
H_2O	water
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
HLB	hydrophilic lipophilic balanced
HIV	human immunodeficiency virus
HMBC	heteronuclear multiple bond coherence (^{1}H-^{13}C)
HPLC	high performance liquid chromatography
HRESIMS	high-resolution electron spray ionisation mass

	spectrometry or spectrum
HSQC	heteronuclear single quantum coherence (^{1}H-^{13}C)
HTS	high throughput screening
Hz	Hertz
IC$_{50}$	hibition concentration at 50%
LC-MS	liquid chromatography/mass spectrometry
LLE	lead-like enhanced
LLEE	lead like enhanced extract
LLEF	lead like enhanced fraction
m	multiplet
Me	methyl
mg	milligram
MHz	mega Hertz
MIC	minimum inhibitory concentration
min	minute(s)
mL	millilitre
MRSA	methicillin-resistant *Staphylococcus aureus*
MS	mass spectrometry
MW	molecular weight
m/z	mass to charge ridio
NMR	nuclear magnetic resonance
NP	natural product
PA	Pseudomonas aeruginosa
PC	principal component
PCA	principal component analysis
ppm	parts per millon
q	quartet
RNA	ribonucleic acid
Ro5	rule-of-five
ROESY	rotational nuclear overhauser effect spectroscopy
s	singlet

SA	*Staphylococcus aureus*
t	triplet
TB	tuberculosis
TCM	Traditional Chinese medicine
TFA	trifluoroacetic acid
UV	ultra-violet
δ	ppm of the applied magnetic field
μg	microgram
μL	microlitre

Acknowledgement

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people.

Firstly, I would like to express my sincere gratitude to my principal supervisor Professor Ronald J Quinn for the continuous support to my Ph.D study, for his patience, motivation, and immense knowledge. His advice on both research as well as on my career have been priceless. And during the most difficult times when writing this thesis, he gave me the moral support and the freedom I needed to move on. I could not have imagined having a better supervisor for my Ph.D study.

I would like to thank my external supervisor Dr. Tanja Grokvic for her insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

My sincere thanks also go to my Chinese supervisors Professor Lixin Zhang and Xueting Liu from Chinese Academy of Sciences who encouraged me to pursue my doctoral degree in Australia and who made it possible for me to obtain a scholarship from Griffith University. They have provided invaluable guidance to me through their knowledge of science and their personal integrity whilst supervising my PhD studies. Many thanks go to them for sharing their wisdom with me.

I would like to thank the present and past members of Quinn group for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last three and a half years. I am especially grateful to Dr. Asmaa Boufridi and Dr. Wendy Loa-Kum-Cheung for their help in the laboratory and research facilities.

Last but not the least, a special thanks to my family. Words cannot express how grateful I am to my parents for always believing in me and encouraging me to follow my dreams and for all of the sacrifices that they've made on my behalf. I would also like to thank all of my friends who supported me in writing, and incented me to strive towards my goal. I would like express appreciation to my boyfriend Jianying Han who was always my support in the moments when there was no one to answer my queries.

ALL PAPERS INCLUDED ARE CO-AUTHORED

Acknowledgement of Papers included in this Thesis

Section 9.1 of the Griffith University Code for the Responsible Conduct of Research ("Criteria for Authorship"), in accordance with Section 5 of the Australian Code for the Responsible Conduct of Research, states:

> To be named as an author, a researcher must have made a substantial scholarly contribution to the creative or scholarly work that constitutes the research output, and be able to take public responsibility for at least that part of the work they contributed. Attribution of authorship depends to some extent on the discipline and publisher policies, but in all cases, authorship must be based on substantial contributions in a combination of one or more of:
>
> - conception and design of the research project
> - analysis and interpretation of research data
> - drafting or making significant parts of the creative or scholarly work or critically revising it so as to contribute significantly to the final output.

Section 9.3 of the Griffith University Code ("Responsibilities of Researchers"), in accordance with Section 5 of the Australian Code, states:

> Researchers are expected to:
>
> - Offer authorship to all people, including research trainees, who meet the criteria for authorship listed above, but only those people.
>
> - accept or decline offers of authorship promptly in writing.
> - Include in the list of authors only those who have accepted authorship
> - Appoint one author to be the executive author to record authorship and manage correspondence about the work with the publisher and other interested parties.
> - Acknowledge all those who have contributed to the research, facilities or materials but who do not qualify as authors, such as research assistants, technical staff, and advisors on cultural or

community knowledge. Obtain written consent to name individuals.

Included in this thesis are papers in Chapters 2, and 4 which are co-authored with other researchers. My contribution to each co-authored paper is outlined at the front of the relevant chapter. The bibliographic details (if published or accepted for publication)/status (if prepared or submitted for publication) for these papers including all authors, are:

(Where a paper(s) has been published or accepted for publication, you must also include a statement regarding the copyright status of the paper(s).)

Chapter 2: **Miaomiao Liu**, Tanja Grkovic, Lixin Zhang, Xueting Liu*, Ronald J Quinn*. A Model to Predict Anti-tuberculosis Activity: Value Proposition for Marine Microorganisms. *J. Antibiot.* 2016, 69(8): 594-599 (DOI:10.1038/ja.2016.87)

Chapter 4: **Miaomiao Liu**, Wael M. Abdel-Mageed, Biao Ren, Wenni He, Pei Huang, Xiaolin Li, Krishna Bolla, Hui Guo, Caixia Chen, Fuhang Song, Huanqin Dai, Ronald J Quinn, Tanja Grkovic, Xiaoping Zhang, Xueting Liu*, Lixin Zhang*. Endophytic *Streptomyces* sp. Y3111 from Traditional Chinese Medicine Produced Antitubercular Pluramycins. *Appl. Microbiol. Biotechnol.* 2013, 98(3): 1077-1085 (DOI: 10.1007/s00253-013-5335-6)

Appropriate acknowledgements of those who contributed to the research but did not qualify as authors are included in each paper.

(Signed) _____ (Date)_____

Miaomiao Liu

(Countersigned) _____ (Date)_____

Supervisor: Ronald J. Quinn

Chapter I Introduction

Abstract: Tuberculosis has been one of the most global harmful infectious diseases. Due to the severe drug resistance and HIV-TB co-infection, it is urgent to search for new anti-TB therapeutics. Microbes have made a phenomenal contribution to the health and wellbeing of people throughout the world. In addition to producing many primary metabolites, they are capable of making secondary metabolites, which constitute half of the pharmaceuticals on the market. This chapter presents the background of this project, including tuberculosis disease, drug resistance, current clinical agents and candidates, anti-TB natural products from marine microbes and endophytes, and three important strategies used in drug discovery, including high throughput screening, one strain-many compounds microbial biosynthetic elicitation strategy, and NMR fingerprinting.

1.1 Tuberculosis

Tuberculosis (TB) has been, and still is one of the world's biggest threats. It causes ill health among millions of people each year and ranks as the leading cause of death from an infectious disease worldwide. It is estimated that one-third of the world's population is infected with the tubercle bacillus.[1] While only a small percentage (5-15%) of infected individuals will develop clinical TB, there were approximately 10.4 million new cases and 1.4 million deaths in 2015.[1] *Mycobacterium tuberculosis* is thus responsible for more human mortality than any other single microbial species.

1.1.1 Pathogen

Mycobacterium tuberculosis, the causative agent of TB is difficult to diagnose in latent form and shows little or no clinical symptoms. It typically affects the lungs (pulmonary TB) but can affect other sites as well (extra-pulmonary TB). Pulmonary TB is by far the most common form of tuberculosis, accounting for 86% of all new and relapsed cases worldwide, based on WHO data. In lung infections, *M. tuberculosis* is typically inhaled into the body through the mouth or nose, and passed through the airways where it reaches the alveolar space in the lungs.[2] In the majority of the cases, the bacteria are able to evade immune system defences to establish a latent infection. It does so by the arrest of phagosome-lysosome fusion, resistance against reactive nitrogen intermediates (RNI) and nitric oxide (NO) and interference with major histocompatibility complex (MHC) Class II antigen-presentation.[3] Reactivation disease occurs when latent bacteria from old, scarred, granulomatous lesions are reactivated into an active, virulent state. Reactivation disease is most frequently triggered when the host immune response weakens or is suppressed. For example, most of the HIV positive individuals are low in $CD4^+$ T counts and face 10% risk per year of TB reactivation disease.[4]

1.1.2 HIV-TB co-infection

The concurrent management of HIV/TB co-infection remains a challenge owing to adverse effects, complex drug interactions, overlapping toxicities and tuberculosis - associated immune reconstitution inflammatory syndrome. An estimated 1.2 million (11 %) of the 10.4 million people who developed TB in 2015 were HIV-positive.[1] The co-evolution of bacteria with their host has led to bidirectional evolutionary pressures, eventually leading to refinement of survival mechanisms in both species that culminate in

symbiotic relationships.[5] In the individual host, the two pathogens, Mycobacterium tuberculosis and the HIV type 1 (HIV), appear to have great impact on each other, accelerating the deterioration of immune functions.[6] HIV-related dysfunction of local immune responses in HIV/TB co-infected individuals reduces the ability of granuloma, an organized structure comprising epithelioid macrophages surrounded by a rim of lymphocytes, to restrain tubercle bacilli, leading to their eventual multiplication and dissemination, thereby culminating in severe pathology.[7] Furthermore, co-infection with *M. tuberculosis* negatively impacts the HIV immune responses and up-regulates HIV replication in blood and pulmonary lymphocytes and alveolar macrophages ex vivo.[8]

1.1.3 Drug resistance

Anti-tuberculosis drug resistance is a major public health problem that threatens progress made in TB care and control worldwide. Drug resistance arises due to improper use of antibiotics in chemotherapy of drug-susceptible TB patients.[9] Because the treatment regimen for TB is long and complex, many patients are unable to complete the course of treatment, enabling their disease to develop drug-resistance. Once a drug-resistant strain has developed, it can be transmitted directly to others.[10]

All TB

TB with any drug resistance

MDR-TB
drug resistance to at least the first line drugs isoniazid and rifampin

XDR-TB
drug resistance to the first line drugs isoniazid and rifampin and to specific second line drugs
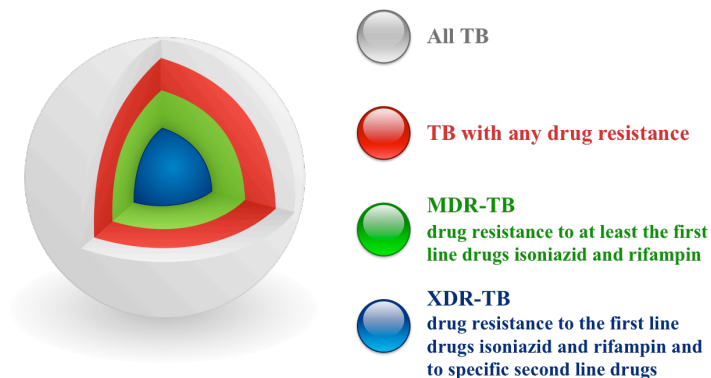
Figure 1.1 The definition of drug resistant tuberculosis.

Much attention has focused on the burden of multi-drug resistant (MDR) TB.[11] MDR-TB is defined as cases of TB that are resistant to at least rifampicin and isoniazid (Figure 1.1). At present, MDR-TB continues to be a significant problem, not only in the developing countries, but also in the Baltic region, parts of the former Soviet Union, and other areas of the world.[12] The emergence of drug-resistant TB is often attributed to a failure to implement adequate control programs for tuberculosis and to appropriately manage cases of the disease. Therefore, resistance to TB accurately reflects the poor

quality of control programs and is a direct consequence of poor therapeutic practices with respect to the use of anti-TB drugs.[13] There were an estimated 480, 000 new cases of MDR-TB worldwide in 2015 and more than half of these cases were in India, China and the Russian Federation.[1]

Extensively drug-resistant tuberculosis (XDR-TB) is a type of multidrug-resistant tuberculosis (MDR-TB) that is resistant to the front-line drugs and two or more of the six classes of second-line drugs (Figure 1.1). Transmission of XDR-TB is in clusters and follows similar transmission patterns as ordinary TB.[14] This makes it difficult to put appropriate barriers to the transmission of the deadly strains. To make matters worse, proper diagnosis involving culture and sensitivity tests is the most commonly used diagnostic method especially in the poor countries. This may take from 6 to 16 weeks before XDR-TB is confirmed during which time it is likely to have spread to other patients.[15] The next challenge is that there are limited treatment options for XDR-TB especially in the developing countries and this makes the disease virtually untreatable and puts HIV positive people particularly at a greater risk.[16]

Three years after XDR-TB was first reported in 2006, Velayati *et al.* drew attention to the emergence of totally drug-resistant (TDR) tuberculosis in a cohort of 15 patients from Iran, resistant to all first- and second-line drugs.[17] The majority of prescriptions were ineffective and would only have served to further amplify resistance, converting MDR tuberculosis to XDR tuberculosis and TDR tuberculosis, which can be identified as a form of incurable tuberculosis.[18]

In summary, drug resistant TB kills nearly 2 million people each year and seriously threatens global TB control.[1] To prevent drug resistant TB from spreading, there is an urgent need for new diagnostic tools and more importantly, novel and more effective anti-TB drugs and vaccines are needed to be developed.

1.2 TB drugs and candidates in clinical trials

A number of efficacious anti-tubercular agents were discovered in the late 1940s and 1950s with the last, rifampin, introduced in the 1960s.[19] However, the emergence of multi-drug resistant strains of *M. tuberculosis* makes it necessary for the discovery of new drugs, and also implements other modalities of treatment. The pipeline of potential new treatments has been fulfilled with 17 compounds in clinical trials or preclinical

development with promising activities against sensitive and resistant *Mycobacterium tuberculosis* strains.[20] The mechanisms of all first-line drugs, second-line drugs and candidates in clinical trials are listed in Figure 1.2.

1.2.1 First-line treatment

Standardized short-course chemotherapy-rifampicin (**1**) and isoniazid (**2**) for 6 months, supplemented with pyrazinamide (**3**) and ethambutol (**4**) in the first 2 months-is effective against drug-susceptible tuberculosis under controlled conditions.[21] Although rates of treatment-limiting side effects vary, mild adverse effects are common.[22] The limitations of control programs are most evident in settings that are frequently characterized by poverty, high levels of HIV co-infection, and poor access to a high standard of treatment.[23] Therefore, short and simple regimens that are effective, safe, and robust during routine programmatic conditions are urgently needed.

Rifampicin (**1**)   Isoniazid (**2**)   Pyrazinamide (**3**)   Ethambutol (**4**)
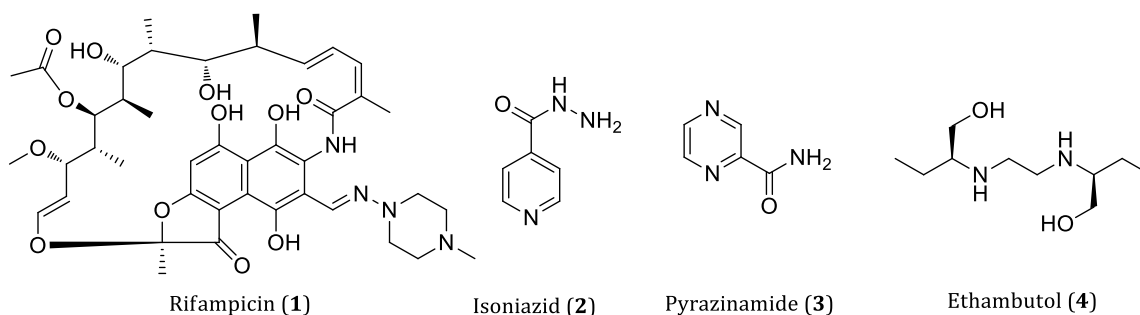
1.2.2 Second-line treatment

Second-line drugs are used for the treatment of MDR and XDR TB. They include aminoglycosides (kanamycin (**5**) and amikacin (**6**)), cycloserine (**7**), terizidone (**8**), ethionamide (**9**), protionamide (**10**), capreomycin (**11**), aminosalicylic acid (**12**), and fluoroquinolones ofloxacin (**13**) and levofloxacin (**14**). Treatment regimens for MDR and XDR tuberculosis are longer, less effective, less tolerable, and more expensive than is standardized short-course chemotherapy, and include the use of injectable drugs, such as kanamycin (**5**) and amikacin (**6**).[24] Containment of the spread of the MDR and XDR tuberculosis will be extremely difficult without treatment regimens that are shorter, safer, more effective, and less expensive than those that are available.[25]
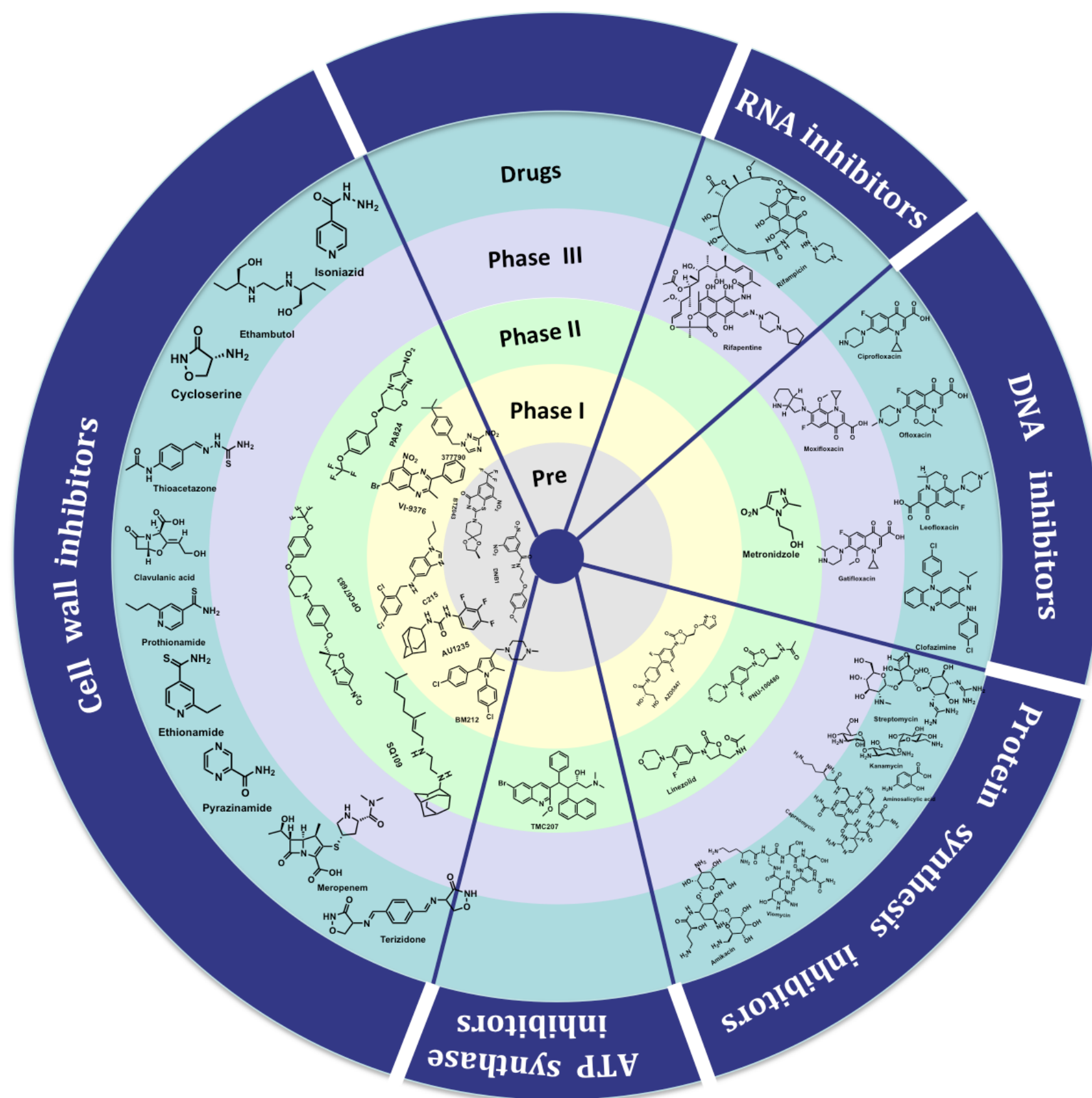
Figure 1.2 Mechanism of action of 39 approved TB drugs and candidates in clinical trials.

Kanamycin (**5**)

Terizidone (**8**)

Amikacin (**6**)

Cycloserine (**7**)

Ethionamide (**9**)

Prothionamide (**10**)

Aminosalicylic acid (**12**)

Ofloxacin (**13**)

Leofloxacin (**14**)

Capreomycin (**11**)
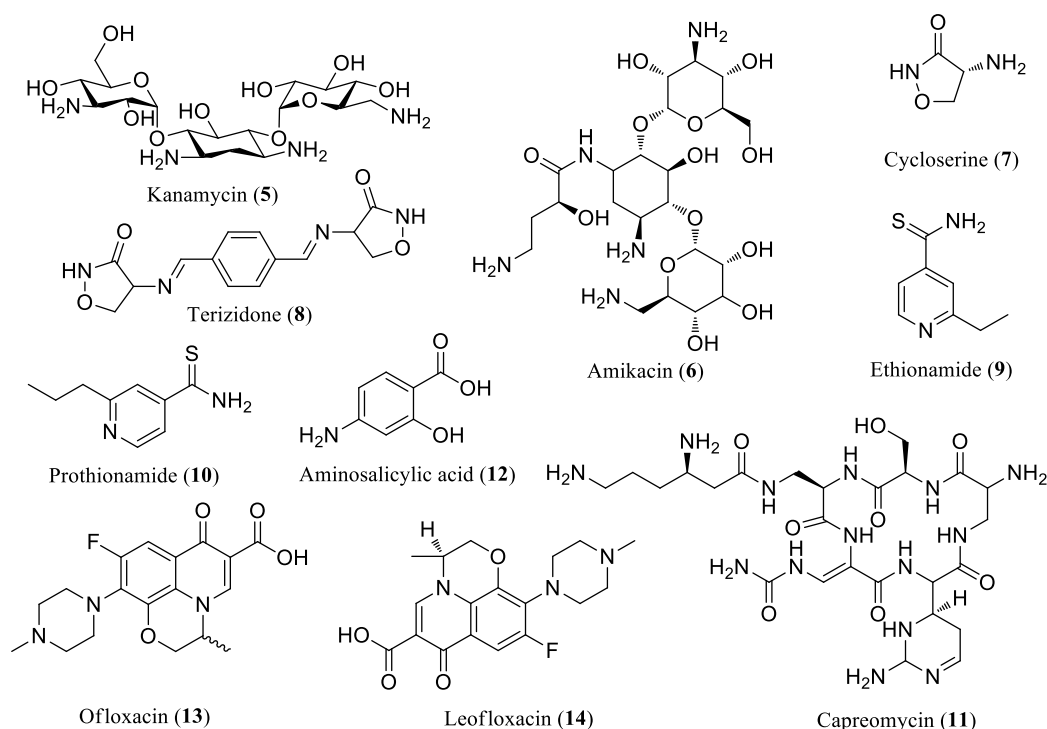
1.2.3 Microbes-derived TB drugs

Among the 11 currently used nature-derived TB drugs, 7 of them were either isolated from microbes or semi-synthesized from microbial natural products (Figure 1.3), suggesting the high potential of microbial compounds to be TB drugs.

Rifampicin (isolated from *Amycolatopsis rifamycinica*)

Capreomycin (isolated from *Streptomyces capreplus*)

Rifapentine (semi=synthesized from rifampicin)

Streptomycin (isolated from *Streptomyces griseus*)

Viomycin (isolated from *Streptomyces puniceus*)

Amikacin (semi-synthesized from streptomycin)

Kanamycin (isolated from *Streptomycin kanamyceticus*)
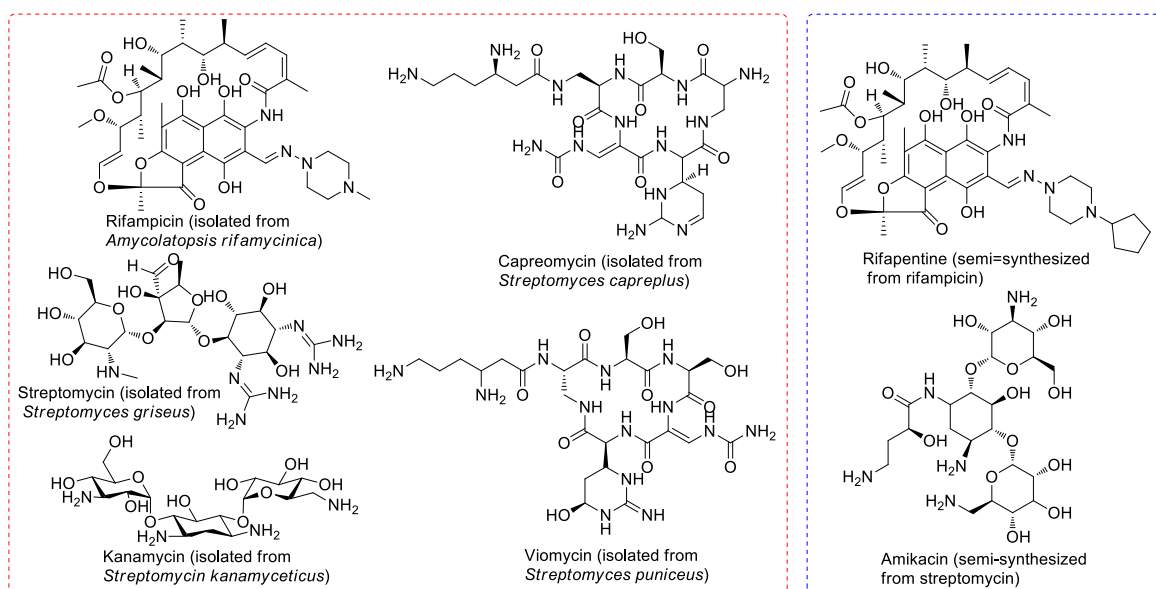
Figure 1.3 Seven TB drugs sourced or based on the NP scaffold isolated from a microbe.

1.2.4 Drugs in clinical development

Eight small molecules in clinical development for TB as well as three existing drugs that are redeveloped or repurposed for TB and five new chemical compounds that are specifically developed for TB are introduced below.

Fluoroquinolones are broad-spectrum antimicrobial drugs that target DNA gyrase.[26] Gatifloxacin (**15**) and moxifloxacin (**16**), the most recently developed fluoroquinolones, have shown better *in vitro* activity against *M. tuberculosis* compared to ofloxacin (**13**) and levofloxacin (**14**), the previous used fluoroquinolones.[27] According to results from a mouse model of tuberculosis infection, moxifloxacin-containing regimens have the potential to shorten treatment of drug-susceptible TB from 6 months to 4 months.[28] Phase 3 trials are in progress for investigation of whether treatment of drug-susceptible tuberculosis can be shortened to 4 months by substitution of gatifloxacin (**15**) for ethambutol (**4**), or moxifloxacin (**16**) for ethambutol (**4**) or isoniazid (**2**).[29]

Nitroimidazoles are antimycobacterial compounds that are equally active against drug-susceptible and drug-resistant TB.[30] Two nitroimidazoles, PA-824 (**17**) and OPC-67683 (**18**), are currently in clinical trials for the treatment of TB and the outcome of these may determine the future directions of drug development for anti-tubercular nitroimidazoles.[31] PA-824 is a prodrug that requires the mycobacterial glucose-6-phosphate dehydrogenase (FDG1) or its cofactor, coenzyme F420, to be transformed into an active form.[32] No cross-resistance between PA-824 with standard anti-TB drugs has been observed.[33] OPC-67683 is a member of the nitroimidazo-oxazole family.[34] Early bactericidal activity and Phase 1 studies have been completed, and the compound is being assessed in a Phase 2 trial for the treatment of MDR tuberculosis.[35]

TMC207 (**19**), the most active diarylquinoline, is highly potent against drug-susceptible and drug-resistant strains of *M tuberculosis*. It has a novel mechanism by inhibiting the mycobacterial ATP synthase enzyme and exhibits synergy with pyrizinamide *in vitro*.[36] Compared with isoniazid or rifampicin, TMC-207 showed no early bactericidal activity for at least the first 4 days, but showed similar activity to rifampicin or isoniazid from 5–7 days when administered at 400 mg per day in mice model. [37] The safety, tolerability, and efficacy of TMC-207 when added to individualised treatment for newly diagnosed MDR tuberculosis is being investigated in a Phase 2 trail.[38]

Oxazolidinones exert their antimicrobial activity by inhibiting protein synthesis by binding to the 70S ribosomal initiation complex.[39] Linezolid (**20**), the only approved drug in this class, showed weak early bactericidal activity against *M. tuberculosis* in patients with cavitary pulmonary TB.[40] Long-term use of linezolid has been associated with

Gatifloxacin (**15**)

Moxifloxacin (**16**)

PA824 (**17**)

OPC-67683 (**18**)

Linezolid (**20**)

PNU-100480 (**21**)

TMC207 (**19**)

SQ109 (**22**)
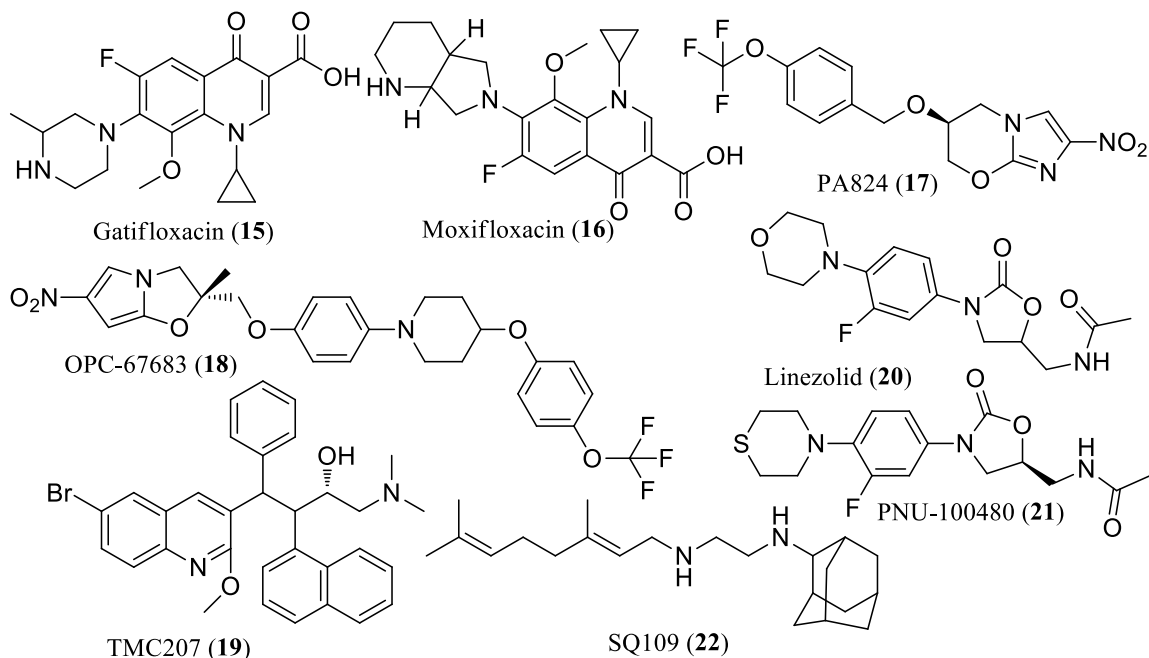
cumulative toxicity, including peripheral and optic neuropathy.[41] PNU-100480 (**21**) is an analogue of linezolid that is being developed. It showed slightly better activity against *M. tuberculosis* in vitro than did linezolid, but substantially improved activity in mouse models of TB.[42] A combination regimen of PNU-100480, moxifloxacin, and pyrazinamide was more active than the standard regimen of rifampicin, isoniazid, and pyrazinamide.[43] These findings suggest that PNU-100480 has the potential to shorten treatment of drug-susceptible and drug-resistant tuberculosis.[44]

SQ-109 (**22**) is a derivative of ethambutol (**4**), but differs in its mode of action.[45] No cross-resistance is seen between SQ109 and the other drugs for TB.[46] It interacts synergistically with isoniazid and rifampicin, even with rifampin-resistant strains.[47] In a mouse model of established *M. tuberculosis* infection, substitution of SQ-109 for ethambutol in the standard regimen improved 15.5% activity of mycobacterial killing.[48]

TB is now found in every corner of the globe and is threatening gains made in TB control. However, it has been nearly 40 years since the introduction of a novel compound (rifampicin) for the treatment of tuberculosis. In the past decade, ten compounds have

progressed into the clinical development pipeline, including six new compounds specifically developed for tuberculosis. Despite this progress, the global drug pipeline for tuberculosis is still insufficient to address the needs of treatment. There are three reasons given for needing new tuberculosis drugs:

(1) to improve current treatment by shortening the total duration of treatment or by providing for more widely spaced intermittent treatment,

(2) to improve the treatment of MDR-TB, and

(3) to provide for more effective treatment of latent tuberculosis infection.[49]

1.3 TB-active natural products isolated from microbes sourced from unique environments

The impact of natural products on human beings has been enormous, and their study continues to influence research in the fields of chemistry, biology, and ecology. While advances in synthetic and combinatorial chemistry have given rise to notable successes in the development of new drugs, the perceived value of natural products has not recessed when it comes to treating infectious diseases. In this part, we review the important role of natural products isolated from unique environment and their activity against TB, and the current status of these natural products as new TB drug leads.

A literature review of the TB-active natural products sourced from the unique environments was conducted. Based on the known knowledge of the literature at the start of the research project, in total 60 compounds were identified from marine microbes and 66 compounds were identified from endophytes by 2014. The chemical properties and biological activities of these potential compounds are discussed below.

1.3.1 Marine environment

Oceans cover 70% of the Earth's surface, and depending on their location, temperature, salinity, a great diversity of beings inhabits them, forming a much more extensive phylogenetic diversity than in the terrestrial environment.[50] The association of a great genetic diversity of marine organisms and a great ecological diversity, ranging from accessible beaches to extreme ocean depths make oceans a unique and rich source for active compounds for the pharmaceutical industry.[51] The physical factors that influence the marine microbes are a) salinity and pH, b) low water potential, c) high concentrations of salts, such as sodium ions, d) low temperature, e) oligotrophic nutrient conditions and

f) high hydrostatic pressure, the last three parameters being unique to the deep-sea environment.[52] In the last 40 years, various active compounds with unique structural and chemical characteristics have been discovered from sponges, ascidians and marine microorganism, which have started the "blue biotechnology" concept for use of these organisms and derivatives to develop new drugs.[53]

1.3.2 Anti-TB natural products from marine microbes

A literature search targeting natural products isolated from marine microbes that have reported anti-mycobacterium activity resulted in 60 molecules. The structures and biological activities of 18 of the active compounds, which were the most active and some of which exhibited rare and novel structures, are discussed below. All of the 60 marine microbial compounds with reported anti-TB activity will be discussed in chapter II.

Abyssomicins are polycyclic polyketides that possesses unprecedent complex structures and impressive biological profiles. Abyssomicins B (**23**), C (**24**) and D (**25**) were first isolated as secondary metabolites of the gram-positive actinomycetete *Verrucosispora* AB-18-032 collected from the bottom (depth 289 m) of the Japanease Sea.[54] A subsequent 2007 reinvestigation of AB-18-032 led to three additional co-metabolites in the form of abyssomicins G (**26**), H (**27**), and atrop-abyssomicin C (**28**).[55] In 2012, three new members of this rare class, abyssomicin J (**29**), K (**30**) and L (**31**) were identified

Abyssomicin B (**23**)

Abyssomicin C (**24**):
atrop-abyssomicin C (**28**):

Abyssomicin D (**25**)

Abyssomicin G (**26**)

Abyssomicin H (**27**)

Abyssomicin J (**29**)

Abyssomicin K (**30**): R=H
Abyssomicin L (**31**): R=CH$_3$
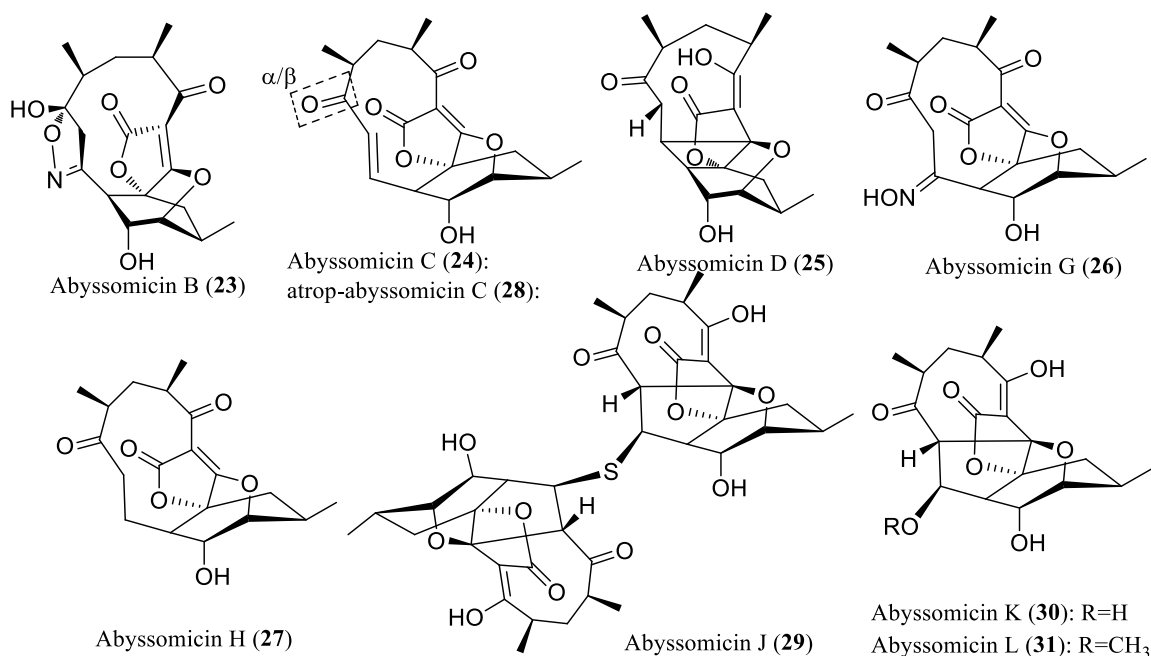
from a marine-derived actinomycetete, *Verrucosispora* sp. (MS100128) collected from

the South China Sea (deep-sea -2733 m).[56] Among all the abyssomicins compounds, abyssomicins J and abyssomicin C showed most significant anti-TB activity, with MIC 2.26 μM and 2.16 μM, respectively.[56] The mechanism of the *M. tuberculosis* inhibition of these compounds is not fully understood, but abyssomicin C was hypothesized to act as an inhibitor of *p*-aminobenzoic acid (pABA) biosynthesis.[57] Detailed analytical studies into abyssomicin Michael addition chemistry demonstrated that abyssomicin J (**29**) can act as a prodrug, thus responding to oxidative activation to selectively deliver the anti-TB antibiotic atrop-abyssomicin C (**28**).[56]

Three aminolipopeptides named trichoderins A (**32**), A1 (**33**) and B (**34**) were isolated from a culture of marine sponge-derived fungus *Trichoderma* sp. 05FI48.[58] The three compounds showed potent anti-mycobacterial activity against *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis* H37Rv under both aerobic conditions and dormancy-inducing hypoxic conditions, with MIC values in the range of 0.1–1.75 μM.[58] These results indicated that trichoderins were effective against *Mycobacterium* sp. in both actively growing and dormant states. Structure-activity relationship analysis suggested that the AHMOD (2-amino-6-hydroxy-4-methyl-8-oxo- decanoic acid) moiety is important for anti-mycobacterial activity of trichoderins.[58] The transformant of *M. smegmatis*, which over-expressed a part of genes that coded mycobacterial ATP synthase, was found to exhibit a resistance to trichoderin A. In addition, trichoderin A reduced ATP contents in *M. bovis* BCG, indicating that the anti-mycobacterial activity of trichoderins comes from the inhibition of ATP synthesis.[59]

Trichoderin A (**32**): R$_1$= R$_2$= CH$_3$

Trichoderin A1 (**33**): R$_1$= R$_2$= CH$_3$

Trichoderin B (**34**): R$_1$= R$_2$= H
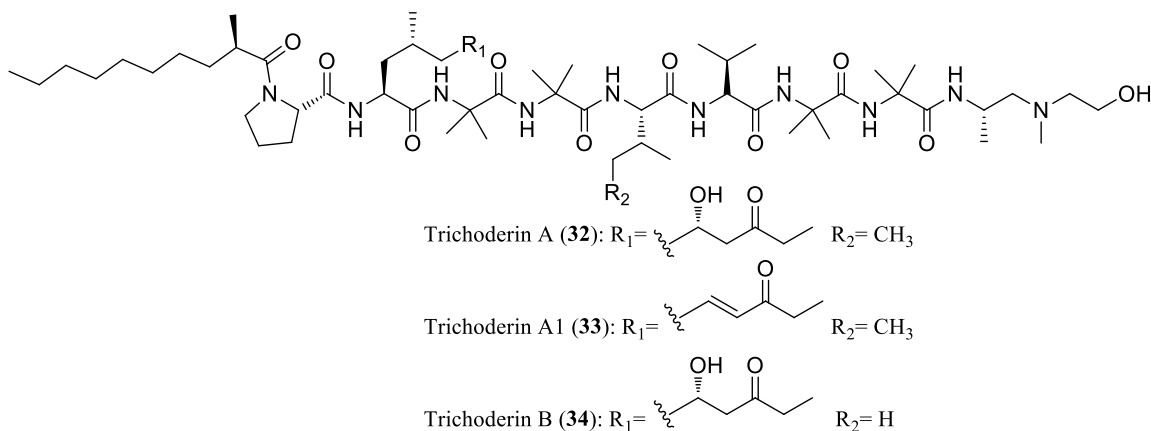
Peniphenones B (**35**) and C (**36**), possessing an unusual benzannulated 6,6-spiroketal, have been isolated from a mangrove fungus *Penicillium dipodomyicola* HN4-3A

collected from the South China Sea.[60] Peniphenones exhibited strong inhibitory activity against MptpB with IC_{50} values of 0.16 ± 0.02 and 1.37 ± 0.05 μM, respectively.[60] The promising inhibitory activities together with their unique structures indicate that they could represent a new type of leads for the development of new drugs tackling TB infections.

Brunsvicamides B (**37**) and C (**38**) were isolated from a sponge-associated cyanobacterium *Tychonema* sp. in 2006.[61] The compounds are characterized by six amino acids, five of which form a 19-membered ring structure, closed by an amide bond between the carboxylic group of the C-terminal Phe and the ε-amino group of the N-terminal D-Lys. Brunvicamides were inactive ($IC_{50} > 100\mu$M) toward five protein phosphatases, including CDC25a, VHR, PTP1b, SHP2, and MptpA. The *Mycobacterium tuberculosis* enzyme MptpB was, however, potently inhibited by brunsvicamide B (IC_{50} 7.3 μM) and brunsvicamide C (IC_{50} 8.0 μM).[61] The fact that the peptides do not interact with other phosphatases illustrates the high selectivity toward the mycobacterium-derived enzyme.

Peniphenone B (**35**)   Peniphenone C (**36**)   Brunsvicamide B (**37**)

Brunsvicamide C (**38**)   8'-*O*-demethylnigerone (**39**)   8'-*O*-demethylisonigerone (**40**)
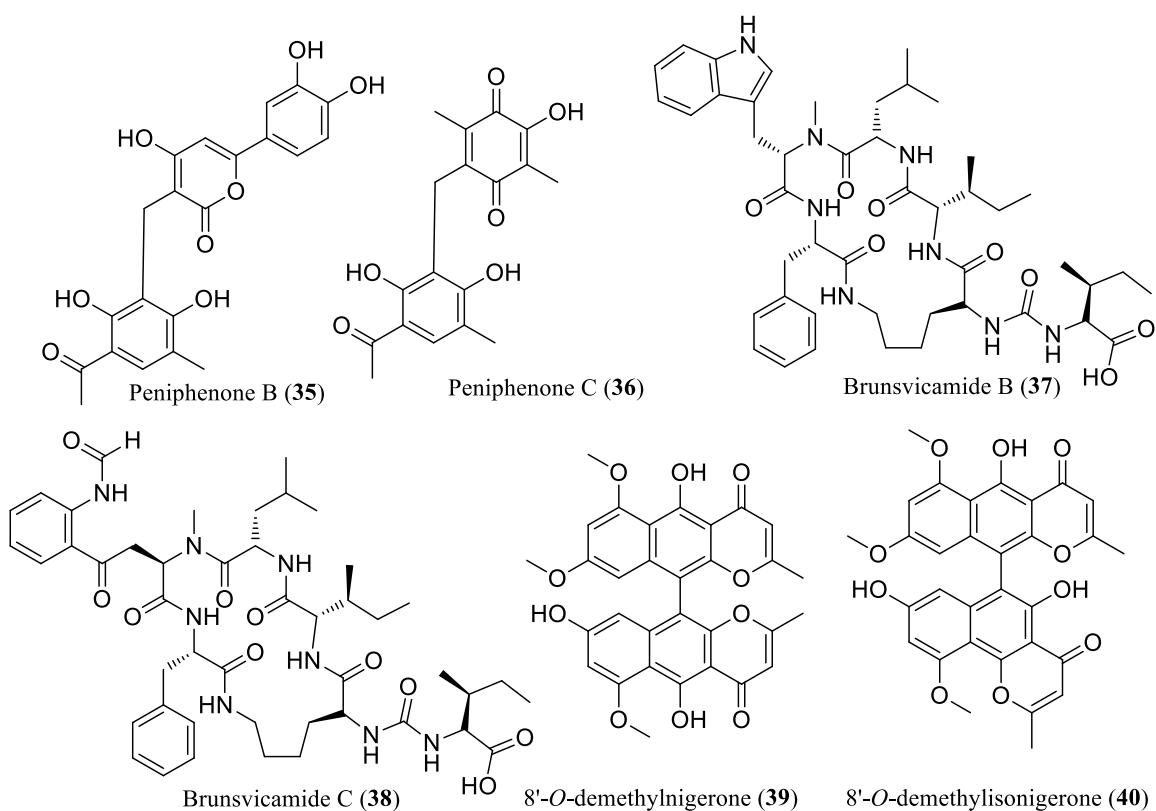
Two dimeric naphtho-γ-pyrones, 8'-*O*-demethylnigerone (**39**) and 8'-*O*-demethylisonigerone (**40**), were isolated from the EtOAc extract of the fungal strain WZ-

551 of *Aspergillus carbonarius*.[62] The compounds demonstrated antimycobacterial activities against *M. tuberculosis* H37Rv, with MIC values of 43.0 and 21.5 μM, respectively. The SAR studies strongly indicated that the presence of conjugated double bonds in the pyrone rings were crucial for activity.[62]

It can be concluded that marine microbes have the ability to produce secondary metabolites unlike those found in any terrestrial species. In summary of anti-TB natural products from the marine environment, 16 of 60 compounds were found to have activity of 10 μM or greater, suggesting marine microbes could become a promising source for anti-TB drug exploration. Research into marine microbes in particular is hampered by the level of knowledge concerning the basic biology and culture techniques for marine-derived organisms, and many more application of methods designed for the isolation, culture and chemical investigation of these organisms are needed.[63]

1.3.3 Endophyte environment

Endophytes are microorganisms (mostly fungi and bacteria) that live in the inter-cellular spaces of plant tissues without causing apparent diseases.[64] Some endophytic microbes have been shown to protect plants from herbivores or to be responsible for the synthesis of novel and useful small molecules that may be involved in a host-endophyte relationship.[65] As a direct result of the role that these secondary metabolites may play in nature, they may ultimately be shown to have applicability in medicine. In fact, a comprehensive study has indicated that 51% of biologically active substances isolated from endophytic fungi were previously unknown, compared to only 38% of novel small molecules from soil microflora.[66] Overall, the rationale for studying endophytic microbes as potential sources of new medicine is related to the fact that this is a relatively unexplored area of biochemical diversity.[67] Furthermore, the search is driven by the fact that the contribution of the endophyte to the plant may be to provide protection to it by virtue of antimicrobial compounds that it produces.[68] Some of these compounds may be of medicinal interest, since they possess antifungal, antibacterial, antimalarial, and a host of other biological activities. Finally, of major concern to the medical community is the latent toxicity of any prospective drug to the higher organisms such as animal and human tissues. Since the plant is also a eukaryotic system, in which the endophyte exists, the antibiotics made by the endophyte may have reduced cell toxicity; otherwise, death of the

host tissue may occur.[69] Thus, the plant itself has naturally served as a selection system for microbes having bioactive molecules with reduced toxicity toward higher organisms.

1.3.4 Anti-TB natural products from endophytes

Up to 2014, 66 natural products identified from endophytes have been found to exhibit TB growth inhibition activity and 14 of the most effective compounds are discussed in detail in the following section, with the entire set analyzed for physico-chemial properties in the following chapter.

Gliotoxin (**41**)   methylthio-gliotoxin (**42**)
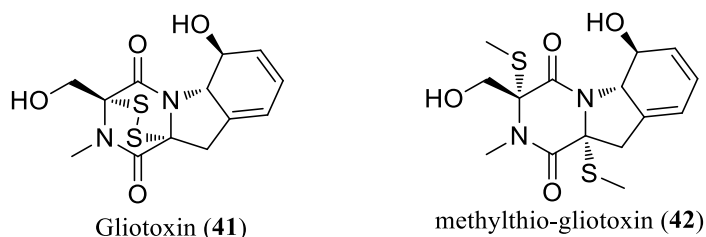
Gliotoxin (**41**), a sulfur-containing diketopiperazine produced by several species of molds, has a number of interesting chemical and biological properties. Purified gliotoxin was found to inhibit growth of human and bovine tubercle bacilli when incorporated in various liquid culture media in concentrations of 0.012 to 0.06 μg/mL.[70] Recently, the chemical investigation of the endophytic *Penicillium* sp. BCC16054 led to the isolation and purification of its methylthio-analog (**42**).[71] Biological assays showed that gliotoxin and methylthio-glitoxin exhibited very strong antitubercular activities with MIC values of 0.382 ng/mL (1.17 nM) and 48.8 ng/ mL (0.14 μM), respectively.[71]
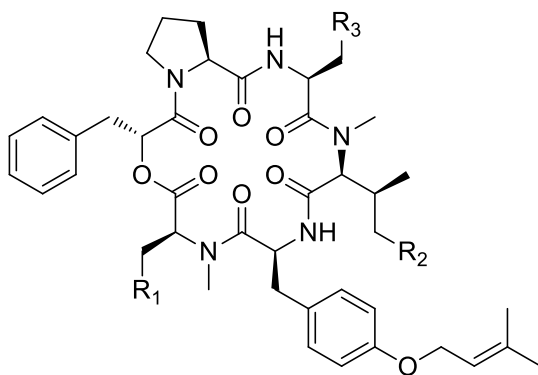
Pullularin A (**43**): R$_1$= H, R$_2$= CH$_3$, R$_3$ = OH
Pullularin B (**44**): R$_1$= CH$_3$, R$_2$= CH$_3$, R$_3$ = OH
Pullularin C (**45**): R$_1$= H, R$_2$= H, R$_3$ = OH
Pullularin D (**46**): R$_1$= H, R$_2$= CH$_3$, R$_3$ = H

Pullularins are cyclohexadepsipeptides that possesses an unusual *O*-prenyl-L- tyrosine

residue and a diverse range of bioactivities, such as anti-malarial, anti-cancer and anti-HSV (Herpes simplex virus).[72] In the search for new types of anti-TB agents, four new pullularins A–D (**43-46**) were isolated from the endophytic fungus *Pullularia* sp. BCC 8613. Pullularins A–C showed moderate antitubercular activity, with MIC 25 μg/mL, 50 μg/mL and 100 μg/mL, respectively.[72]
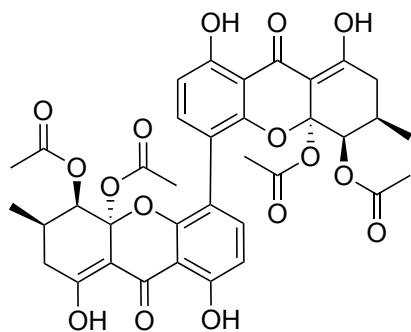
Phomoxanthone A (**47**)
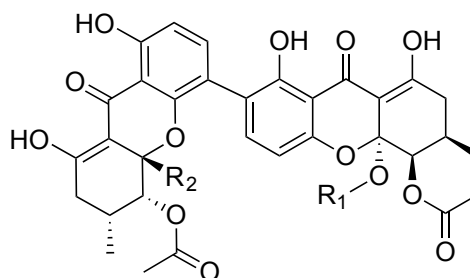
Phomoxanthone B (**48**): R$_1$ = COCH$_3$, R$_2$ =COO CH$_3$
deacetyl-phomoxanthone B (**49**): R$_1$ = H, R$_2$ = CH$_2$OH

The genus *Phomopsis* has been known to be a rich source of bioactive secondary metabolites of diverse structures.[73] Two novel xanthone dimers, phomoxanthones A (**47**) and B (**48**), were isolated from the endophytic fungus *Phomopsis* sp. BCC1323.[74] They exhibited significant activity against tuberculosis (H37Rv strain), with MIC 0.5 μg/mL and 6.25 μg/mL, respectively.[74] Chemical investigation of another endophytic fungus *Phomopsis* sp. PSU-D15 led to isolation of another new member of this unique class, deacetyl-phomoxanthone B (**49**).[75] However, the data suggests that this new compound might display much weaker antimycobacterial activity than phomoxanthones B.[75]
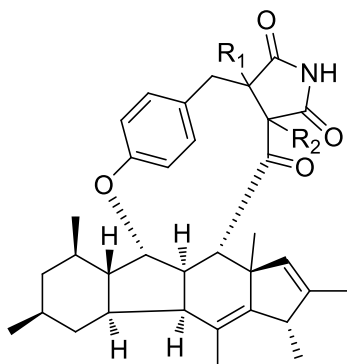
Phomapyrrolidone A (**50**): R$_1$ = H, R$_2$ = H
Phomapyrrolidone B (**51**): R$_1$ = H, R$_2$ = H
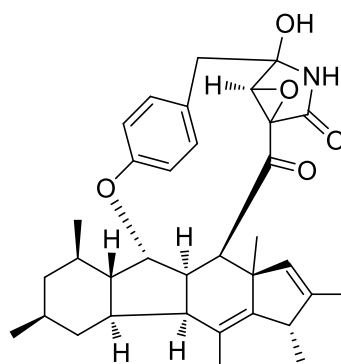
Phomapyrrolidone C (**52**)

Three alkaloids, phomapyrrolidones A−C (**50-52**), bearing a cyclopenta[*b*]fluorene ring system, were isolated from the mycelium extract of the endophytic fungal strain *Phoma*

sp. NRRL 46751, inhabiting *Saurauia scaberrinae*.[76] Phomapyrrolidones B and C have been found to exhibit potent activity against replicating *M. tuberculosis* H37Rv at the concentrations of 5.9 μg/mL and 5.2 μg/mL, while phomapyrrolidone A showed a much weaker *M. tuberculosis* inhibition ability at 20.1 μg/mL.[76] Furthermore, all the three compounds displayed mild activity against nonreplicating H37Rv at MIC values of 41.1 μg/mL, 13.4 μg/mL and 15.4 μg/mL, respectively.[76]

Two new azaphilone derivatives, namely biscogniazaphilones A (**53**) and B(**54**), were isolated from the endophytic fungus *Biscogniauxia formosana* BCRC 33718, derived from the bark of medicinal plant *Cinnamomum* species.[77] Results of the antimycobacterial activities indicated that biscogniazaphilones A and B exhibited potent antimycobacterial activities against *M. tuberculosis* strain H37Rv *in vitro*, showing MIC values of ≤5.12 and ≤2.52 μg/mL, respectively, compared to the clinical drug, ethambutol (MIC 6.25 μg/ mL).[77] Biscogniazaphilone B, with one γ-lactone group between C-6a/C-9, was 2-fold stronger compared to biscogniazaphilone A, indicating that the presence of the γ -lactone group of the azaphilone analogues possibly plays an important role in their antimycobacterial activity.[77]
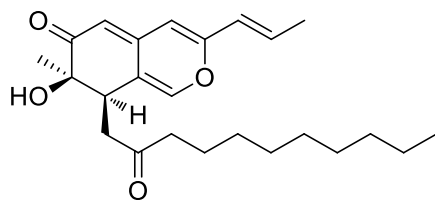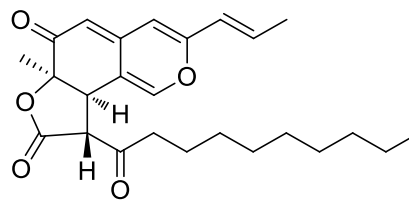
Biscogniazaphilones A (**53**)   Biscogniazaphilones B (**54**)

Endophytes are a relatively poorly investigated source of bioactive and chemically novel compounds. This is witnessed by the discovery of a wide range of anti-TB natural products as mentioned in this chapter. In total, 14 of 66 compounds were found to have activity of 10 μM or greater and some of them had unique mode of action of *M. tuberculosis* inhibition. The high chemical diversity and TB activity make natural products isolated from endophytes become potential TB drug candidates.

Finding a cure for tuberculosis is one of the greatest actual challenges for pharmacology and medicine. There is an extensive research effort aimed at obtaining sufficient compounds of natural origin. As has been demonstrated in this chapter, natural products derived from marine microbes and endophytes have the potential to be sources and leads

of TB drugs. In conclusion, literature review focused on the anti-mycobacterium natural products sourced from extreme environments identified 60 compounds from marine microbes and 66 compounds from endophytes. In this chapter, we outlined several extreme active compounds from these two environments and discussed their chemical and biological characteristics.

In responding to the challenge of TB drug discovery and with the consideration of the potential of active metabolites from marine microbes and endophytes, this thesis aims to: (1) conduct a physico-chemial analysis of anti-TB natural products from marine microbes and endophytes; (2) run a in vitro anti-BCG screen of marine microbial extracts and endophytes extracts; (3) isolate and identify anti-TB natural products from the hits of the screening assay; (4) use one strain many compounds strategy and NMR fingerprint technology to identify unique anti-TB metabolites that are only produced under specific conditions. A brief overview of the general experimental procedures is outlined below.

1.4 Important strategies of identifying active microbial natural products in drug discovery
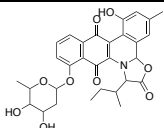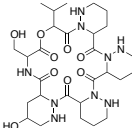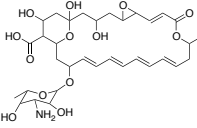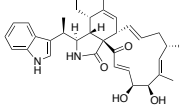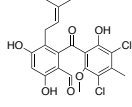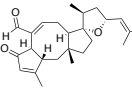
1.4.1 One strain many compounds (OSMAC) strategy

While many organisms produce secondary metabolites, the actinomycetetes, in particular the bacterial genus *Streptomyces*, are an especially rich source.[78] The view in the past was that the screening conducted previously had yielded most of the chemical diversity that exists in nature and so finding new metabolites might be difficult.[79] However, the sequencing of streptomycete genomes demonstrated that every streptomycete genome has the genetic capacity to generate 20 – 40 distinct secondary metabolites, far more than had been predicted previously.[80] The difficulty is that many of these gene clusters are expressed at low levels or only expressed under very specific conditions during growth in the laboratory. This has led to efforts aimed at activating these genes and identifying previously unreported molecules.[81-83]

OSMAC (one strain - many compounds) is a strategy that was developed in the early 2000s by Zeeck and co-workers to maximize the productivity of a single microbe through turning on silent or cryptic biosynthetic genes by varying accessible cultivation parameters, such as media composition, pH, temperature, oxygen supply, quality and quantity of light, bioreactor platform, or the addition of precursors or enzyme inhibitors.[84]

Some examples of active microbial natural products with the use of OSMAC strategy are documented in Table 1.1.

Table 1.1 Examples of active microbial natural products identified through the application of the OSMAC strategy

Compound name	Structure	OSMAC strategy	Biological activity	Sources	Year
Jadomycin B		Heat shock	Antibacteria	*Streptomyces venezuelae*	1992[85]
Piperidamycin		Addition of antibiotics	Antibacteria	*Streptomyces* sp.	2009[86]
Pimaricin		Addition of inducers	Antifungal	*Streptomyces natalensis*	2004[87]
Chaetoglobosins		Addition of inhibitors	Cytotoxic	*Phomospis asparagi*	2005[88]
Pestalone		Co-culture	Antibacteria	*Pestalotia* sp.	2001[89]
3-Anhydro-6-hydroxy-ophiobolin A		Media variation	Antibacteria	*Bipolaris oryzae*	2013[90]

Analysis of metabolites in each fraction is key to targeting the unique compounds only produced by the microbes under specific OSMAC conditions. NMR has a unique place not only in structure elucidation, and characterization of molecules but is now also considered as a major tool in metabolomics studies. Moreover, NMR has also been widely used in combination with different multivariate data analyses methods to do metabolic profiling of various samples.[91]

1.4.2 NGS in drug discovery

Next-generation sequencing (NGS) is a high-throughput methodology that enables rapid sequencing of the base pairs in DNA or RNA samples. It supports a broad range of applications, including gene expression profiling, chromosome counting, detection of

epigenetic changes, and molecular analysis.[92] These next generation sequencing-related methods has been used to deal with complex biological problems, for example, the identi cation of all sequence changes in drug resistant *Mycobacteria tuberculosis*.[93] In addition, NGS technology also creates a tremendous impact in medical science by sequencing of a person's genome sequence of an individual human and the initiation of the personal Human Genome Project.[94]

1.4.3 MS and NMR fingerprinting in drug discovery

Mass spectrometry has been involved into all phases of drug discovery. MS data can give molecular weight, fragmentation and molecular formula that are essential to assign the molecular structure of an unknown compound. Furthermore, the advances in the ionization techniques available for MS combined with other developments, such as the introduction of triple-quadrupole, time-of-flight, and high-mass-resolution instruments (e.g. fourier transform ion cyclotron resonance), have enabled a serial of MS-based tools to address complex problems within biology.[95] For example, native MS has become a fast, simple, highly sensitive and automatable technique with well-established utility for fragment-based drug discovery (FBDD).[96] Native MS is able to directly detect weak ligand binding to proteins, to determine stoichiometry, relative or absolute binding affinities and specificities.[97] Native MS can also be used to target ligand-binding sites, to elucidate mechanisms of cooperativity and to study the thermodynamics of binding.[98, 99]

While mass spectrometry is more sensitive, it does not intrinsically provide enough structural information about the class of the small molecules of the crude extracts or fractions. Metabolite fingerprinting by NMR is a fast, convenient, and effective tool for discriminating between groups of related samples and it could identify the most important regions of the spectrum for further analysis. NMR fingerprinting and multivariate analysis of ^{1}H NMR spectra could be used to compare the overall metabolic composition of microbial extracts, and to assess the impact of culture conditions on the microbial metabolome. These profiles may also be used to compare groups of samples, and significant differences in metabolites provide the basis for novel natural products of the groups.

1.4.4 Data analysis

Analysis of the large amount of data generated using each of different techniques, such as

MS or NMR, and the often subtle differences between samples require the use of statistical programs. Chemometric analysis is therefore a very important feature of modern analytical approaches for the characterization of complex matrices. The most popular chemometric analysis model principal component analysis (PCA) was chosen as suitable analysis tool in this project for the task of comparing different OSMAC conditions and experiments.

Principal component analysis (PCA) is the most popular multivariate statistical technique and it is used by almost all scientific fields. Normally, PCA analyzes a data table representing observations described by several dependent variables, which are, in general inter-correlated.[100] Its goal is to reduce a data table to its most important components by removing correlated characteristics and in order to achieve this goal, PCA computes new variables called principal components which are obtained as linear combinations of the original variables. [101] The first principal component is required to have the largest possible variance and therefore this component will explain the largest part of the inertia of the data table. The second component is computed under the constraint of being orthogonal to the first component and to have the largest possible inertia. [102]
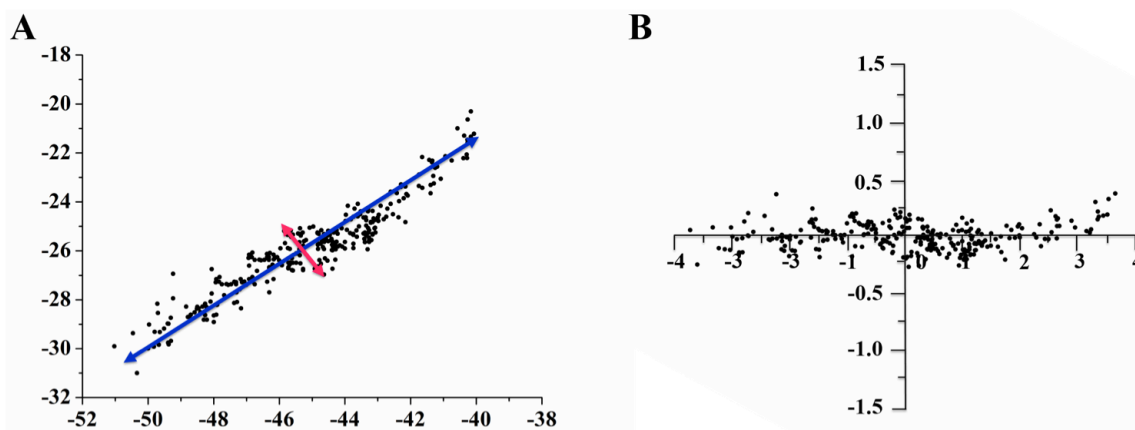
Figure 1.4 Principal component analysis. A. Plot of a random gaussian distribution. The first principal component was directed with blue arrow, and the second principal component was directed with magenta arrow. B. Plot of the random gaussian distribution with two principal components.

The first step to find the components of a principal component analysis is centering the variables then plotting them against each other. Secondly, find the main direction (or the first component) of the points such that we have the minimum of the sum of the squared distances from the points to the component. Add a second component orthogonal to the

first such that the sum of the squared distances is minimum. Lastly, when the components have been found, rotate the figure in order to position the first component horizontally (and the second component vertically).[103] A simple example is given in Figure 1.4.

In the following chapter, we will examine the physico-chemial properties of the TB-active natural products and drugs. With the latest increase in tuberculosis research, our purpose is to facilitate discussion on the issue that compounds from marine microbes and endophytes as being one of the most potential sources for further drugs and therapeutic use in the treatment of TB. Also, we will give a robust method for TB activity prediction of any inconclusive compounds from marine microbes.

Research projects are presented and discussed within chapters 3 to 7 (Figure 1.5). First of all, a cell-based anti-TB screening of microbial crude extracts and pure compounds were examined (chapter 3). One of the hits found from chapter 3 was scaled-up and six compounds, including 4 new compounds were identified by the use of bioassay-guided isolation (chapter 4). To overcome the shortcomings of bioassay-guided isolation, a NMR fingerprint-guided isolation was conducted on 50 small-scaled microbial crude extracts (chapter 5). With the consideration of new structure as well as anti-TB activity, 32 OSMAC conditions were applied on 13 actinomycetes (chapter 6) and based on NMR fingerprints and data analysis results, target compounds in outlier fractions were isolated and evaluated (chapter 7).
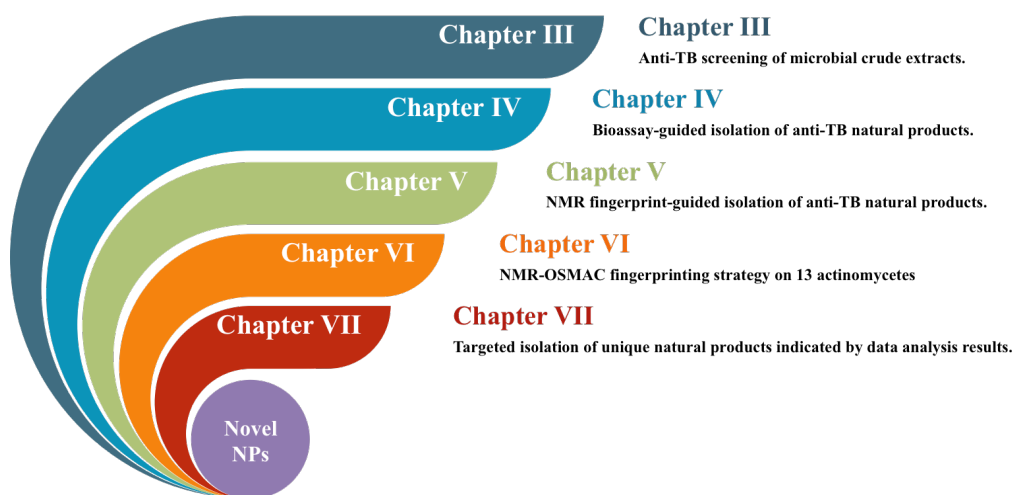
Figure 1.5 A general overview of chapter 3 to chapter 7.

Chapter II Physico-chemical analysis of TB-active NPs sourced from marine microbes and endophytes

<u>STATEMENT OF CONTRIBUTION TO CO-AUTHORED PUBLISHED PAPER</u>

This chapter includes a co-authored paper. The bibliographic details of the co-authored paper, including all authors, are:

Miaomiao Liu, Tanja Grkovic, Lixin Zhang, Xueting Liu*, Ronald J Quinn*.

My contribution to the paper involved:

Literature research, data calculation and analysis using Instant J chem and ChemGPS, the development of the prediction model and manuscript writing.

(Signed) _____ (Date)_____
Miaomiao Liu

(Countersigned) _____ (Date)_____
Corresponding author of paper: Ronald J Quinn / Xueting Liu

(Countersigned) _____ (Date)_____
Supervisor: Ronald J Quinn

Abstract: The development of new antibiotics effective against all strains of tuberculosis (TB) is needed. To evaluate the potential of marine microbe-derived and endophyte natural products as anti-TB leads, we analyzed and compared the physico-chemial properties of 39 current TB drugs and candidates against 60 confirmed marine microbial and 66 endophyte-derived TB-active natural products. We showed that anti-TB natural products sourced from marine microbes and endophytes have a large overlap with TB drug-like space. A model to predict potential anti-TB drugs is proposed.

2.1 Introduction

TB has several barriers to drugs, such as delivery to lung, granuloma, penetration of the unusual cell wall of *M. tuberculosis* (mycolic acid). Modern drug discovery concentrates on understanding of the relationship between physico-chemial properties and potency. The primary objective of the following analysis is to evaluate whether natural products from marine microbes or endophytes could be potential leads of TB drugs. Moreover, is there something else could we learn from the physico-chemial properties?

2.1.1 Rule of five (Ro5) and physico-chemial properties

The drug-likeness of every compound is associated with its acceptable aqueous solubility and intestinal permeability. By calculating the physico-chemial properties, it's possible to predict the oral bioavailability of each candidate. Ro5 considers orally active compounds and defines four simple physico-chemial parameter ranges (MW$\leq$ 500, log P $\leq$ 5, H-bond donors $\leq$ 5, H- bond acceptors $\leq$ 10) associated with 90% of orally active drugs that have achieved Phase II clinical status (Figure 2.1).[104] If a compound fails the Ro5 there is a high probability that oral activity problems will be encountered. However, passing the Ro5 is no guarantee that a compound is drug-like.[105] TPSA is an additional descriptor that has been shown to correlate with passive molecular transport through membranes, and therefore allows prediction of transport properties of drugs. It was shown that PSA $\leq$ 140 Å and number of rotatable bonds $\leq$ 10 is an efficient and selective criterion for a drug-like molecule.[106]
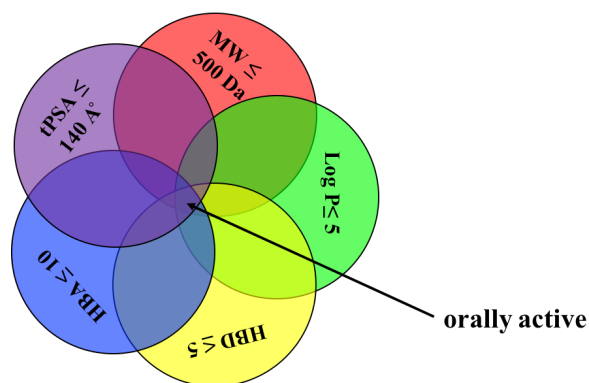
Figure 2.1 Illustration of physico-chemial properties associated with an active drug.

2.1.2 Overview of Chem-GPS

In the pharmaceutical industry for development of new leads, internet technology offers an exceptional framework to develop public scientific applications. The internet has

Table 2.1 35 ChemGPS-NP descriptors

Number	Abbreviation	Description
1	MW	molecular weight
2	Sv	sum of atomic van der Waals volumes (scaled on C atom)
3	Se	sum of atomic Sanderson electro- negativities (scaled on C atom)
4	Sp	sum of atomic polarizabilites (scaled on C atom)
5	Mv	mean atomic van der Waals volume (scaled on C atom)
6	Me	mean atomic Sanderson electro- negativity (scaled on C atom)
7	nAT	number of atoms
8	nSK	number of non-hydrogen atoms
9	nBT	number of bonds
10	nBO	number of non-hydrogen bonds
11	nBM	number of multiple bonds
12	ARR	aromatic ratio
13	nCIC	number of rings
14	RBN	number of rotatable bonds
15	RBF	rotatable bond fraction
16	nDB	number of double bonds
17	nAB	number of aromatic bonds
18	nC	number of carbon atoms
19	nN	number of nitrogen atoms
20	nO	number of oxygen atoms
21	nX	number of halogens
22	nBnz	number of benzene-like rings
23	nCar	number of aromatic carbon atoms (sp2)
24	n_amid	number of amides
25	nROH	number of aliphatic hydroxy groups
26	nArOH	number of aromatic hydroxy groups
27	nHDon	number of donor atoms for hydrogen bonds (N and O)
28	nHAcc	number of acceptor atoms for hydrogen bonds (N, O, and F)
29	Ui	unsaturation index
30	Hy	hydrophilic factor
31	AMR	Ghose-Crippen molar refractivity
32	TPSA(NO)	topological polar surface area using N and O
33	TPSA(Tot)	topological polar surface area using N, O, S, and P
34	ALOGP	Ghose-Crippen octanol-water partition coefficient
35	LAI	Lipinski alert index (drug-like index)

become a central source for information, tools, and services to support medicinal chemists and drug discovery researchers and, over the last years, a growing number of

web-based tools for data analysis in chemistry have been made available.[107, 108] One such tool is ChemGPS-NP, which is a principal component analysis (PCA)-based global chemical positioning system tuned for exploration of biologically relevant chemical space, i.e. those areas of chemical space most likely to include biologically active compounds.[109]

In ChemGPS-NP, the selected main rules include aspects of size, shape, lipophilicity, polarity, polarizability, flexibility, rigidity and hydrogen bond capacity.[110] The ChemGPS-NP space map coordinates are t-scores from PCA using a carefully selected subset of 35 descriptors (Table 2.1) that evaluate the above-mentioned rules on a total set of 1779 chosen satellite and core structures.[111] The 35 descriptors were then explained by 8 respective principal components that could be mapped onto a consistent eight-dimensional map. The four most significant PCs explain 77% of the variance and can be interpreted as follows: PC1 represents size, shape and polarizability, PC2 corresponds to aromatic- and conjugation-related properties, PC3 describes lipophilicity, polarity and H-bond capacity, and PC4 expresses flexibility and rigidity.[112] Any compound with a known chemical structure can be positioned onto this map using interpolation in terms of PCA score prediction. From the results, the properties of the compounds can be compared and easily interpreted together with trends and clusters.

2.1.3 Datasets in this study

In the present study five datasets of structures were used as a basis to compare physico-chemical and structural features. These included the following (Figure 2.2):

(1) 39 approved TB drugs and candidates,

(2) 60 TB-active natural products sourced from marine microbial strains identified in the literature review,

(3) 66 TB-active endophytes identified in the literature review,

(4) 157 natural products, including 48 TB-active and 109 TB-inactive compounds, isolated from Prof. Lixin Zhang's lab in IMCAS,

(5) 336 TB-inconclusive natural products from marine microbes identified from literature.

Figure 2.2 Information of 5 datasets in this chapter. N: natural products; S: synthetic compounds; Compound sources are presented in different colors, blue: various sources; green: marine microbes; red: endophytes.

All the compounds, imported as SMILES files from Chemdraw, were positioned in chemical space according to their physico-chemial properties predicted by ChemGPS-NP (Figure 2.3).

Figure 2.3 Schematic diagrams summarizing the process of ChemGPS analysis of the compounds.

2.2 Physico-chemial properties of TB drugs and microbial compounds

Herein, we calculated the MW, log P, HBA, HBD, tPSA and number of rotatable bonds of 39 TB drugs, 66 TB-active NPs from endophytes, 60 TB-active NPs and 336 TB-untested NPs isolated from marine microbes by importing their SMILES files into the program Instant JChem [Instant JChem 3.0.4, 2009 ChemAxon Ltd. (http://www.chemaxon.com)]. The calculated data for the 60 mycobacteria-active natural products and 66 TB-active NPs from endophytes are shown in Table 2.2 and 2.3 and the percentages of the 4 data sets compliant with Lipinski's rule of five are depicted in Figure 2.4.

It was found that around 77% of TB drugs comply with all of the Lipinski's parameters or have just one violation (Figure 2.4). An even higher proportion (89%) of active NPs from endophytes has complied with the Ro5. However, the marine microbial mycobacteria-active natural products data set only had 53% of compounds that were compliant. Unlike mycobacteria-active natural products, 83% of the compounds in the TB-untested dataset exhibited full compliance or just one violation with Ro5.

Table 2.2 Physico-chemial properties of 60 *M. tb*-active natural products from marine microbes.

Cpd no.	Name	MW	LogP	H bond acceptors	H bond donors	TPSA	Rotatable bonds	No. violations	Activity (μM)
1	Pitipeptolides A	808.02	5.07	7	3	180.52	10	2	247.9
2	Pitipeptolides B	810.03	5.57	7	3	180.52	11	2	247.2
3	Massetolide A	1,140.41	1.54	15	13	386.19	28	3	4.39
4	Viscosin	1,126.38	1.1	15	13	386.19	27	3	8.89
5	Cyclomarin A	1,043.30	3.37	11	7	253.27	15	3	191.9
6	Sufactin	894.15	4.01	10	8	267.3	15	2	224
7	Sansanmycin	866.98	-5.08	13	10	290.85	18	3	11.55
8	Sansanmycin B	332.44	-1.94	8	5	124.02	6	0	24.1
9	Sansanmycin C	332.44	-1.94	8	5	124.02	6	0	60.2
10	Caprazamycin B	1,146.28	-0.47	21	7	360.71	33	3	2.73
11	Brunsvicamides B	859.07	4.49	8	8	230.93	13	2	7.3
12	Brunsvicamides C	891.06	3.69	10	8	261.31	15	2	8
13	Trichoderin A	1,163.58	4.29	13	10	313.88	38	3	0.1
14	Trichoderin A1	1,145.56	5.52	12	9	293.65	37	4	1.75
15	Trichoderin B	1,149.55	3.85	13	10	313.88	37	3	0.11
16	Beauvercin	783.95	7.27	6	0	139.83	9	2	5.1
17	Avermectin B1a	873.08	5.85	13	3	170.06	8	3	18.3
18	Valinomycin	1,111.32	5.92	12	6	332.4	9	4	0.9
19	Pitiprolamide	905.13	4.65	8	2	192.04	7	1	221.2

20	Peniphenones B	398.36	3.13	7	5	144.52	4	0	0.16
21	Peniphenones C	330.33	3.38	6	3	111.9	3	0	1.37
22	Nanaomycin *β*A	288.3	1.37	5	2	83.83	2	0	27.8
23	Nanaomycin *α*A	316.31	1.86	5	1	89.9	3	0	25.3
24	Mollemycin A	1,315.46	-0.98	25	10	413.53	18	3	3.2
25	Scytoscalarol	415.66	4.35	4	4	82.13	2	0	110
26	Brevianamide S	678.82	4.84	4	4	113.33	7	1	9.2
27	Ascochitine	276.28	1.36	5	2	83.83	3	0	724.6
28	Ambigol A	484.97	8.14	2	2	49.69	2	1	133.1
29	3-Nitropropionic Acid	119.08	-0.26	4	1	80.44	3	0	3.3
30	hydroxyhalorosellinia A	340.33	-0.83	8	6	147.68	1	1	39
31	Nigrosporin B	304.29	0.71	6	3	104.06	1	0	41
32	Anhydrofusarubin	288.25	1.86	6	2	93.06	1	0	87
33	(+)-aS-alterporriol C	618.54	1.87	13	7	228.35	2	3	8.7
34	Lyngbic acid	256.38	4.26	3	1	46.53	12	0	48.8
35	Lobophorin A	1,157.39	3.13	18	7	291.94	12	3	27.7
36	Lobophorin B	1,187.37	6.05	19	6	309.06	13	4	27
37	Lobophorin G	1,199.42	3.58	18	6	298.01	14	3	13.4
38	Urdamycinone E	532.56	-0.64	10	6	181.82	2	2	5.88
39	Urdamycinone G	514.54	0.63	9	5	161.59	2	1	24.3
40	Dehydroxyaquayamycin	434.44	3.44	7	4	124.29	1	0	14.4
41	Urdamycin E	890.99	1.31	17	7	257.43	8	3	14
42	Actinomycin D	1,255.42	-0.1	16	5	355.54	8	2	6.38
43	Actinomycin X2	1,269.40	-0.76	17	5	372.61	8	2	0.79
44	Abyssomicin C	346.37	1.92	5	1	89.9	0	0	18.1
45	Atrop-abyssomicin C	346.37	1.92	5	1	89.9	0	0	578
46	Abyssomicin H	348.39	1.92	5	1	89.9	0	0	71.8
47	Abyssomicin K	364.39	0.45	6	3	113.29	0	0	68.7
48	Abyssomicin L	378.42	1.09	6	2	102.29	1	0	264.6
49	Abyssomicin J	726.83	2.74	10	4	186.12	2	1	4.3
50	Sydowiols A	384.38	5	5	5	119.61	4	0	36.5
51	Sydowiols B	384.38	5	5	5	119.61	4	0	520.8
52	Sydowiols C	384.38	5	5	5	119.61	4	0	62.5
53	Ophiobolin K	384.55	4.4	3	1	54.37	4	0	166.7
54	6-epi-ophiobolin K	384.55	4.4	3	1	54.37	4	0	677.1
55	6-epi-ophiobolin G	366.54	5.52	2	0	34.14	4	0	185.8
56	4-Deoxybostrycin	320.29	1.05	7	4	124.29	1	0	46.9
57	Nigrosporin	304.29	0.71	6	3	104.06	1	0	65.8
58	demethylnigerone	556.52	5.31	10	3	140.98	3	2	43
59	demethylisonigerone	556.52	5.31	10	3	140.98	3	2	21.5
60	Cyclomarin C	1,027.30	4.26	10	7	240.74	15	2	194.9

Table 2.3 Physico-chemial properties of 66 mycobacteria-active natural products from endophytes.

Cpd no.	Name	MW	LogP	H bond acceptors	H bond donors	TPSA	Rotatable bonds	No. violations	Activity (μM)
1	Asperterpenoid A	386.58	4.55	3	2	57.53	3	0	2.2
2	biscogniazaphilones A	386.53	4.76	4	1	63.60	11	0	13.26
3	biscogniazaphilones B	412.53	5.16	4	0	69.67	10	1	6.12
4	methyldopamine	343.38	2.64	5	3	88.02	7	0	36.4
5	trimethoxyflavone	328.32	2.84	6	1	74.22	4	0	76.2
6	4-methoxycinnamaldehyde	162.19	1.82	2	0	26.30	3	0	259.9
7	enedioxycinnamate	222.24	2.37	3	0	44.76	4	0	262.2
8	cinnamic acid	178.19	1.98	3	1	46.53	3	0	280.9
9	Javanicin	290.27	2.27	6	2	100.90	3	0	86.2
10	3-O-methylfusarubin	336.30	1.32	8	3	122.52	2	0	148.8
11	diastereomer	292.29	1.18	6	2	93.06	3	0	85.6
12	methyl ether	324.33	1.03	7	3	105.45	2	0	154.3
13	Phomoenamide	284.36	0.09	4	4	98.66	7	0	22
14	Pullularin A	775.94	3.03	8	3	174.89	10	1	32.26
15	Pullularin B	789.97	3.55	8	3	174.89	11	1	126.7
16	Pullularin C	761.92	2.59	8	3	174.89	9	1	65.7
17	Phomoxanthones A	722.65	3.93	12	4	238.72	9	2	0.69
18	Phomoxanthones B	722.65	3.93	12	4	238.72	9	2	8.65
19	Agonodepside A	382.46	7.25	4	3	86.99	5	1	75
20	Chaetomanone	588.52	3.24	10	3	190.80	3	1	216.6
21	echinulin	461.65	5.74	2	3	73.99	8	1	169.9
22	2-hydroxymethyl	126.16	0.40	2	1	37.30	1	0	1587.3
23	asterric acid	348.31	3.53	6	3	122.52	6	0	574.7
24	dinitrophenylhydrazone	308.29	3.24	7	2	130.90	5	0	649.4
25	3-nitropropionic acid	119.08	-0.26	4	1	80.44	3	0	3.3
26	(+)-abscisic acid	264.32	2.09	4	2	74.60	3	0	757.6
27	diaportheone A	218.21	1.44	4	2	66.76	0	0	100.9
28	diaportheone B	220.22	1.45	4	2	66.76	0	0	3.5
29	4-Deoxybostrycin	320.30	1.05	7	4	124.29	1	0	46.9
30	Nigrosporin	304.30	0.71	6	3	104.06	1	0	65.8
31	Fusaric Acid	179.22	1.14	3	1	50.19	4	0	111.7
32	Piperine	285.34	2.78	3	0	38.77	3	0	6.1
33	hydronaphthalenone 1	324.33	0.09	7	5	127.45	1	0	77.2
34	hydronaphthalenone 2	340.33	-0.83	8	6	147.68	1	1	36.8
35	desoxybostrycin	308.29	0.40	7	4	116.45	1	0	162.3
36	(+)-aS-alterporriol C	618.55	1.87	13	7	228.35	2	3	8.7

37	pestalols B	316.44	5.04	3	2	57.53	8	1	632.9
38	Phomapyrrolidones A	527.71	5.43	4	1	72.47	0	2	38.1
39	Phomapyrrolidones B	527.71	5.43	4	1	72.47	0	2	11.2
40	Phomapyrrolidones C	543.70	5.02	5	2	88.16	0	2	9.6
41	Ilanefuranone	184.24	1.14	3	1	46.53	4	0	18.04
42	Ilanepyrrolal	245.28	2.19	3	2	62.32	6	0	76.7
43	dihydroxynaphthalen	178.19	0.75	3	2	57.53	0	0	15.4
44	dimethylocta	314.38	3.38	3	1	55.76	6	0	19.5
45	Eriodictyol	288.26	2.53	6	4	107.22	1	0	88.9
46	cis-Dihydroquercetin	304.25	1.82	7	5	127.45	1	0	111.2
47	Syringic acid	182.18	1.07	4	1	55.76	3	0	494.5
48	Methyl syringate	196.20	0.91	4	1	55.76	3	0	229.6
49	(þ)-Glaberide I	280.28	0.71	5	1	74.22	3	0	185.7
50	Methyl linoleate	280.45	6.12	1	0	26.30	14	1	214.3
51	phomoarcherins B	384.47	4.34	4	1	72.83	0	0	130.2
52	enylisochroman-1-one	304.39	5.62	3	2	66.76	6	1	82.2
53	enylisochroman-1-one	302.37	5.26	3	2	66.76	5	1	165.6
54	clicdecahydrofluorene	501.67	5.47	4	1	72.47	1	2	96.5
55	penicolinates A	412.53	6.01	4	0	78.38	15	1	30.3
56	penicolinates b	398.50	4.34	5	1	89.38	14	0	62.8
57	penicolinates C	384.48	5.12	4	0	78.38	13	1	16.3
58	penicolinates D	321.46	4.88	3	1	59.42	13	0	623
59	penicolinates E	307.43	4.67	3	0	48.42	13	0	651
60	phenopyrrozin	215.25	1.26	2	1	40.54	1	0	0.06
61	p-hydroxyphenopyrrozin	215.25	1.26	2	1	40.54	1	0	7.26
62	gliotoxin	326.39	-0.02	4	2	81.08	1	0	0.0012
63	bisdethiobis gliotoxin	356.46	0.61	4	2	81.08	3	0	0.14
64	heraclemycin A	388.42	5.54	5	1	80.67	3	1	257.7
65	heraclemycin C	366.41	6.91	5	2	91.67	5	1	17.1
66	heraclemycin D	364.40	6.55	5	2	91.67	4	1	68.7

The molecular weights histogram of the 4 different datasets shows that the majority of compounds are distributed between 300 Da and 400 Da. About 80% of TB drugs are distributed between molecular weights of 100-500 Da, but only 10% of these are nature-derived drugs (Figure 2.5A). 83% of TB-active NPs from endophytes had molecular weight less than 500 Da. However, only 53% of mycobacteria-active natural products from marine microbes fall into this range. Among the higher molecular weight

compounds, clusters were observed around 800 Da and 1000 Da, exemplified by pitipeptolides A and trichoderin A with Mw of 808.02 and 1163.58 Da, respectively. However, of the 336 TB-untested natural products from marine microbes, 248 compounds (74%) had molecular weights less than 500 Da.
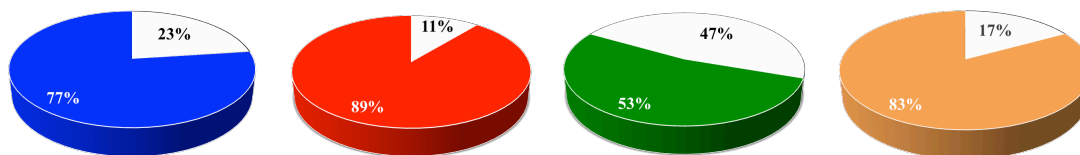
Figure 2.4 Pie chart presentation of the percentage of compounds obeying Ro5. TB drugs (left), TB-active natural products from endophytes (middle left), TB-active natural products from marine microbes (middle right) and untested natural products from marine microbes (right) following or violating Lipinski's rule of five. Non-compliant (more than one violation) is shown in white and compliant (less than two violations) in blue, red, green and orange, respectively.

The distribution of the calculated log P (Figure 2.5B) showed different patterns among the three datasets. The majority of TB drugs were distributed between log P of -2 to 5 and showed a bimodal distribution with peaks in the range of -1 to 1 and 3 to 5. Very similarly, TB-active endophyte NPs also showed a bimodal pattern with log P values from 1-2 and 5-6. The histogram of log P of mycobacteria-active natural products from marine microbes followed a trimodal distribution and showed peaks at 0, 2 and 5. In contrast, in our previous analysis, the calculated log P values of 949 mycobacteria-active natural products showed a normal distribution with the maximum at 3.[113] In order to reach the target, a molecule should have a relatively low polarity (high log P value) to pass through the cell wall of mycobacteria, which mainly consists of mycolic acids.[114] A significant proportion of mycobacteria-active natural products (83%) have a log P between -1 and 5, indicating the high possibility of those mycobacteria-active compounds to cross the lipid barrier in mycobacteria.

The distribution of the calculated HBD (Figure 2.5C) and HBA (Figure 2.5D) showed a wide range of variability. The majority of TB drugs showed a normal distribution for HBD with a maximum at 1 and a bimodal distribution for HBA with maxima at 4 and 7. Mycobacteria-active natural products from marine microbes present a different distribution pattern, showing multiple peaks along the abscissa. Only a relative small percentage of TB drugs (18%) had HBD higher than 5 and HBA higher than 10, while

more than 30% of mycobacteria-active natural products from marine microbes were found to violate these cutoff numbers. Impressively, almost all of the TB-active NPs from endophytes complied HBD (97%) and HBA (95%) limits.

The distributions of the four datasets were similar within the range of tPSA ≤ 140 Å, showing maximum values at 75, 75, 85 and 85 for TB drugs, active endophytes NPs, active marine microbial natural products, and TB-untested, respectively (Figure 2.5E). However, the proportions of compounds with a tPSA value larger than 140 Å showed large differences. Approximately the similar proportions of active endophytes compounds (12.12%), untested compounds (19.64%) and TB drugs (17.94%) had a tPSA value larger than 140 Å, while more than a half of 60 mycobacteria-active natural products (53.33%) had values greater than 140 Å. While tPSA is a descriptor that has been shown to correlate well with passive molecular transport through membranes and, therefore, allows prediction of transport properties of drugs,[115] pulmonary drugs tend to have higher tPSA because pulmonary permeability is less sensitive to polar hydrogen-bonding functionality.[116]
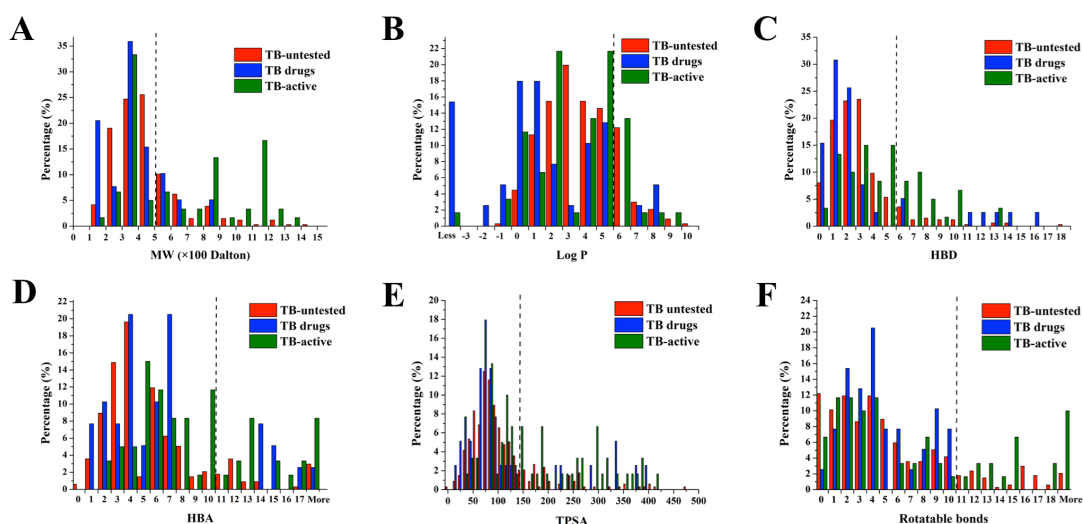
Figure 2.5 Analysis of physico-chemical properties of the compounds. TB drugs (blue), mycobacteria-active natural products from endophytes (red), mycobacteria-active natural products from marine microbes (olive) and TB-untested natural products from marine microbes (magenta). A. molecular weight, B. log P, C. HBD, D. HBA. E. topological surface area, F. rotatable bonds.

Most of the compounds in the four datasets showed less than 5 rotatable bonds (Figure 2.5F). None of 39 TB drugs violated this parameter, whereas 13.98% of untested compounds 12.12% of active endophytes NPs and 30% of mycobacteria-active marine

microbial compounds had greater than 10 rotatable bonds. Greater than 10 rotatable bonds correlates with decreased rat oral bioavailability, although the rotatable bond count does not correlate with in vivo clearance rate.[117]

Trichoderin A: R= CH$_3$
Trichoderin B: R= H

Peniphenone B   Peniphenone C
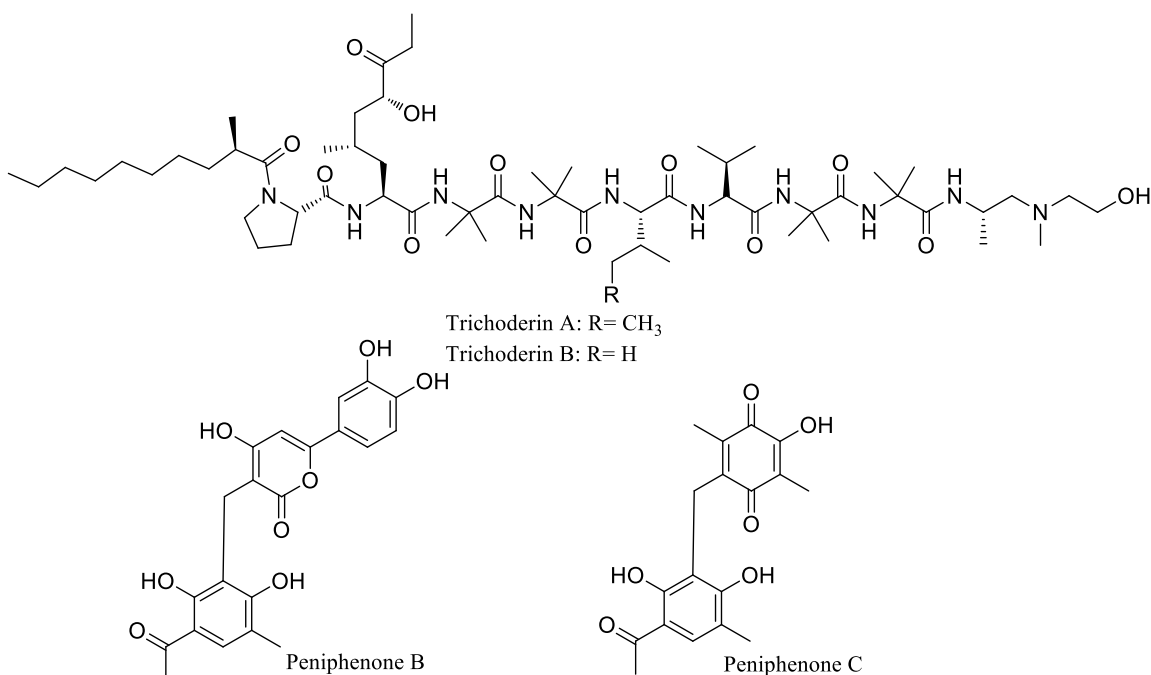
Figure 2.6 Examples of mycobacteria-active natural products isolated from marine microbes.

Interestingly, some active natural products with 3 or even 4 violations showed high anti-TB activity in in vitro experiments. For example, two aminolipopeptides, trichoderins A and B (Figure 2.6), isolated from a culture of marine sponge-derived fungus *Trichoderma* sp. 05FI48, both show 3 violations, Mw (1163.58 and 1149.55), HBA (313.88 and 313.88) and HBD (13 and 13).[58] These two compounds showed potent anti-mycobacterial activity against *M. smegmatis* (0.1 μg/mL and 0.63 μg/mL), *M. bovis* BCG (0.02 μg/mL and 0.02 μg/mL) and *M. tuberculosis* H37Rv (0.12 μg/mL and 0.13 μg/mL) under both aerobic conditions and dormancy-inducing hypoxic conditions, with MIC value of 0.1 nM.[59] In addition, trichoderin A reduced ATP content in *M. bovis* BCG, indicating that the anti-mycobacterial activity of trichoderins comes from the inhibition of ATP synthesis. TMC 207, the only TB drug with ATP inhibitor activity, also had 2 violations, Mw (555.51) and log P (7.13). On the other hand, the analysis identified several potential mycobacteria-active compounds fully compliant to Ro5 as well as significant anti-TB activity. Peniphenones A and B (Figure 2.5), possessing an unusual benzannulated 6,6-spiroketal, isolated from a mangrove fungus *Penicillium dipodomyicola* HN4-3A,

exhibited strong inhibitory activity against MptpB with IC_{50} values of 0.16 ± 0.02 and 1.37 ± 0.05 μM, respectively.[60] The promising inhibitory activities together with their druggable physico-chemial properties indicate that they could represent a new type of lead compound.

2.3 Comparative study of TB-active natural products

2.3.1 TB-active NPs from marine microbes versus TB-active NPs from endophyte

It was demonstrated that marine secondary metabolites can differ fundamentally from their terrestrial counterparts, with more complex chemical structures and different functions.[118] Indeed, this seems to be a reasonable assumption in light of several profound differences between these two environments, i.e. the availability of elements (in particular bromine, chlorine and iodine) in the sea.[119] While the ChemGPS-NP analysis allows a more thorough analysis by the discovery of that a large portion of marine vs endophyte-sourced compounds show a similar occupation pattern in the chemical space (Figure 2.7). This demonstrates that natural products are, at least partially, similar on land and in the sea. This result is not unexpected since there are biosynthetic pathways i.e. terpenoids, acetogenins, aromatic compounds and in some cases also alkaloids, in marine and terrestrial organisms that are responsible for the production of secondary metabolites.[120] These facts may also have contributed to the large overlap in the prediction score plots of the present study. Surprisingly, marine and endophyte-sourced natural products also show similar lipophilicity (Figure 2.7).

Contrary to expectations that marine, compared with endophyte-sourced compounds might be more lipophilic to be effective as chemical defense in the aquatic environment (hydrophilic substances are rapidly dissolved).[121] However, in the PC1 prediction plot (Figure 2.7A) comparing size, clear differences between the marine and endophytes-sourced NPs were detected. Although there was a large portion with an overlap in the smaller size part of the plot, marine substances were more strongly distributed in the large molecular weight part of the plot. This finding demonstrates impressively that marine; in contrast to endophytes substances tend to possess higher molecular weights. Beside the increased molecular weight, the plot also revealed less aromatic structures in marine compounds, which is in agreement with suggestions by Grabowski and Schneider who

stated that marine natural products contain chiral centers and have less aromatic atoms on average.[122]
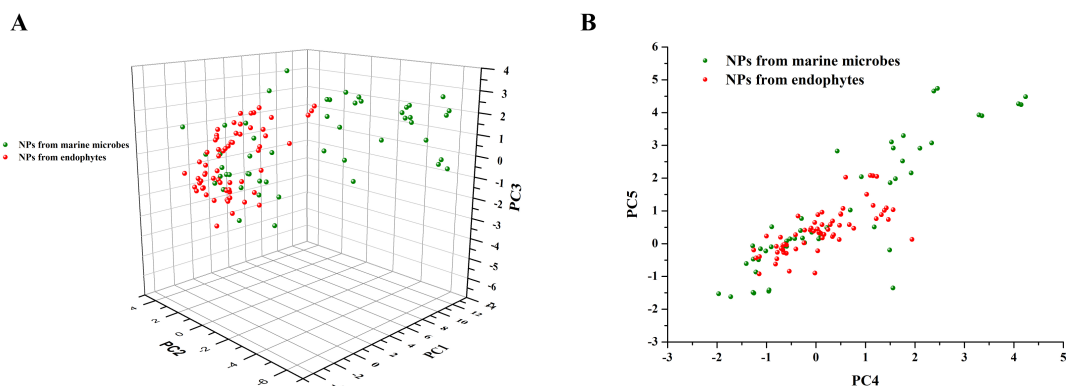
Figure 2.7 Score plot of TB-active natural products from marine microbes and endophytes. A. PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) and B. PC4 (molecular flexibility) versus PC5 (represents mainly the number of halogens, nitrogens).

Furthermore, the prediction plot of PC4/PC5 （Figure 2.7B), demonstrates that, despite an overlap in the central part of the plot, marine compounds stretch out in the upper right corner (large PC4 and PC5 values), while many endophyte compounds are situated in the lower left region (relative small PC4 and PC5 values). This finding clearly indicates that the marine-sourced microbial NPs are more flexible than the endophyte-sourced compounds. This is also evident from the examination of the structures since less aromatic structures contribute to the flexibility of a molecule. This result is also of particular interest, since the main loadings for PC5 include the number of halogens, nitrogens and amides, and the mean atomic Sanderson electronegativities. In other words, marine versus endophyte-sourced compounds contain more halogens, nitrogens and amides.

Typical TB-active marine-sourced molecules are complex, rich in halogens and nitrogens, and contain structures such as long, flexible, aliphatic side chains and flexible polyether structures. TB-active compounds from endophytes, on the other hand, are generally characterized by higher rigidity and less incorporation of halogens and nitrogens in comparison with marine compounds. Moreover, representative endophyte compounds contain larger conjugated systems, including aromatic rings, which enhance the rigidity of the molecules.

2.3.2 TB drugs and candidates versus natural products (marine microbes and endophyte set combined)

Thirty-nine TB drugs and candidates were easily divided into two groups according to their biological origin: 11 nature-derived compounds and 28 synthetic compounds. Interestingly, these two groups showed significant differences after the ChemGPS analysis (Figure 2.8A). Nature-derived TB drugs occupy a very broad range of physicochemical space while most of the synthetic TB drugs distributed in the PC2 positive direction area (high aromaticity). Furthermore, the nature-derived drugs and candidates were divided into three sub-groups in the PC1/PC2/PC3 plot. The first subgroup consists of the injectable drugs, kanamycin, amikacin, streptomycin, capreomycin and viomycin, which showed high molecular weight, the least aromaticity and the lowest lipophilicity. The second subgroup is representative of low molecular weight but high aromaticity, including cycloserine, p-aminosalicyclic acid, pyrazinamide and isoniazid. The last two rifamycin compounds, rifapentine and rifampicin, belong to the third subgroup, possessing the highest molecular weight and high lipophilicity.
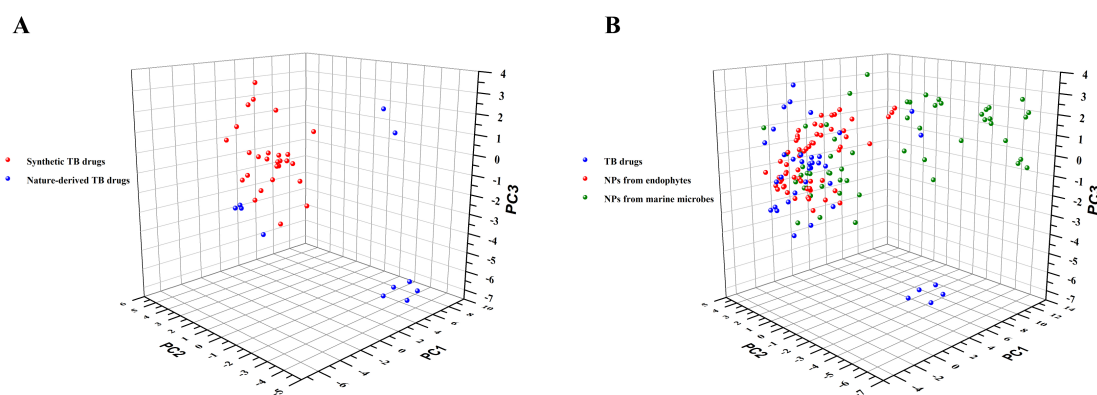
Figure 2.8 Score plot of TB-active natural products and TB drugs. A PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 28 synthetic TB drugs and candidates (red) and 11 nature-derived TB drugs and candidates (blue) and B PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 39 TB drugs and candidates (blue) and 66 natural products from endophytes (red) and 60 natural products from marine microbes (green).

When applying the TB-active natural products from marine microbes and endophytes to this TB drugs and candidates model, they are largely overlapping, at least in the first three dimensions (Figure 2.8B), strongly arguing that natural products have the potential to

serve as an important source for TB drugs. However, the compounds on the right hand, which have higher molecular size, relative low aromaticity and rather non-polar, do not coincide with any of the known drugs and candidates, indicating that these regions have yet to be investigated in TB drug discovery. These compounds include massetolide, viscosin and trichoderins (**32-34**). Based on the well-known argument in medicinal chemistry that compounds with similar structures and properties often have similar biological mechanism, this information indicates a possible new mode of action for the *M. tuberculosis* inhibition.

2.3.3 Regions of the chemical space

In the above section, natural products from marine microbes and endophytes were found to cover unique spaces that lack representation in TB medicinal chemistry compounds, indicating these two sources have the potential to be investigated for anti-TB drugs. Thus, the TB-active natural products from these two special environments were subsequently analyzed by dividing them into different regions in chemical space. Each of the regions was analyzed in terms of occupancy with regard to both chemical properties and biological activities of the compounds.

The whole chemical space of each dataset has been divided into 27 different regions according to the values of the first three most significant PCs (explaining 71% of the variance). The scores specification of each region is listed in Figure 2.9. Some regions had low density for the simple reason that their location implies an impossible combination of properties; i.e. there are limits for individual properties, and a compound cannot simultaneously be small and highly lipophilic and have several H-bond donors and acceptors.

The basic interpretation of the 27 regions of ChemGPS-NP can be as follows: the size of compounds increases in the positive direction of PC1 (i.e. small molecules: regions 1, 4, 7, 10, 13, 16, 19, 22 and 25; medium size molecules: regions 2, 5, 8, 11, 14, 17, 20, 23 and 26; large molecules: regions 3, 6, 9, 12, 15, 18, 21, 24 and 27); compounds are increasingly aromatic in the positive direction of PC2 (i.e. less aromatic molecules: regions 1-9; medium aromatic molecules: regions 10-18; more aromatic molecules: regions 19-27); lipophilic compounds are situated in the positive direction of PC3 while predominantly polar compounds are located in the negative PC3 direction (i.e. polar

molecules: regions 1, 2, 3, 10, 11, 12, 19, 20 and 21; medium-polar molecules: regions 4, 5, 6, 13, 14, 15, 22, 23 and 24; non-polar molecules: regions 7, 8, 9, 16, 17, 18, 25, 26 and 27).

Marine microbial NPs				Endophytel NPs		
PC1	PC2	PC3		PC1	PC2	PC3
-4 ~ 2	-5 ~ -2	-3 ~ -1		-5 ~ -1	-2 ~ 0	-3 ~ -1
2 ~ 8	-2 ~ 1	-1 ~ 1		-1 ~ 3	0 ~ 2	-1 ~ 1
8 ~ 14	1 ~ 3	1 ~ 3		3 ~ 7	2 ~ 4	1 ~ 3
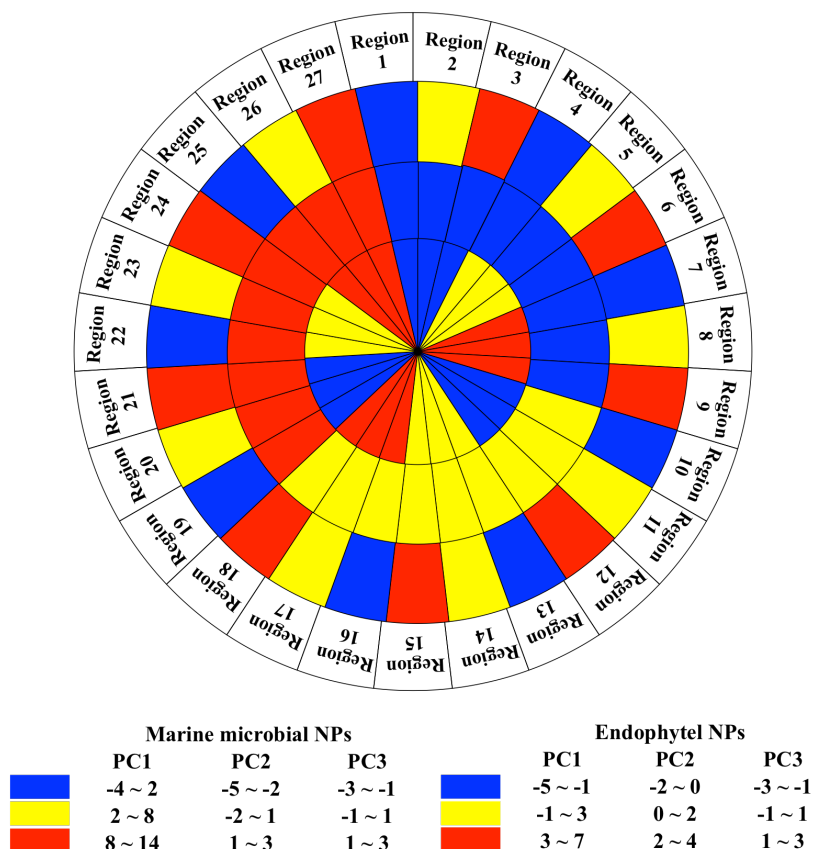
Figure 2.9 The scores specification of each region. PC1: the outer circle, PC2: the middle circle, PC3: the inner circle.

2.3.4 TB-active natural products isolated from marine microbes

All of the marine TB-active microbial natural products were compared with the corresponding TB drugs in the same region. Structures of active NPs examples and drugs were presented in Figure 2.10.

Region 13, representative of low molecular weight, moderate polarity and aromaticity, holds the most active natural products (12) and TB drugs (9) that belong to two different classes distinguished by the mechanisms of DNA inhibition and protein synthesis inhibition （Figure 2.11). The 27 regions can be classified into three classes according to the proportion of TB drugs and mycobacteria-active natural products in each region.
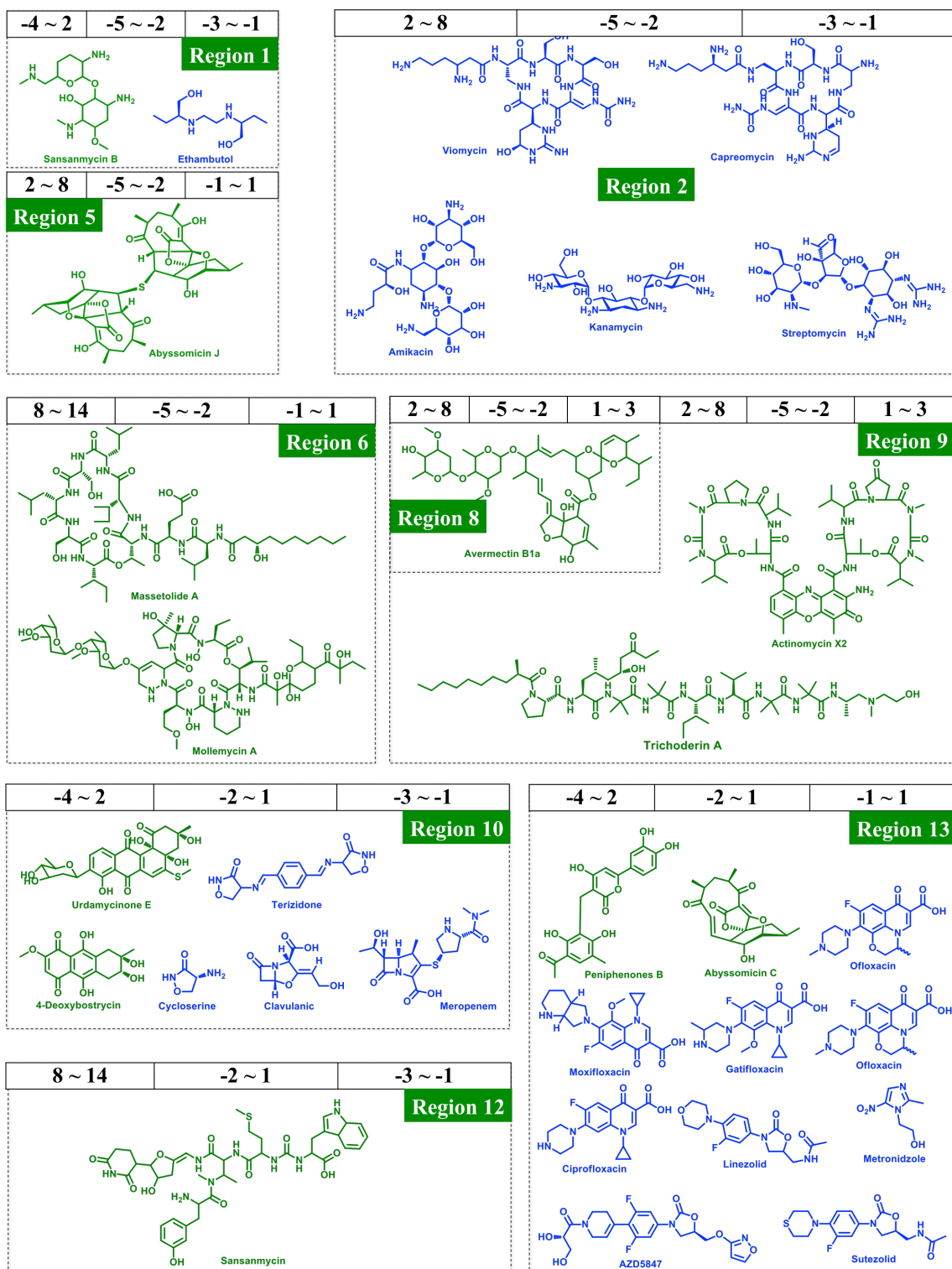
Figure 2.10 Score specifiction of the marine microbes regions with typical structure examples. TB drugs and candidates are in blue and active marine microbial NPs are in green, PC1, PC2 and PC3 are shown on the top of each region.

2 ~ 8	-2 ~ 1	-1 ~ 1

Region 14

Urdamycin E

8 ~ 14	-2 ~ 1	1 ~ 3

Region 18

Cyclomarin A Rifapentine Rifampicin

8 ~ 14	-2 ~ 1	-1 ~ 1

Region 15

Brunsvicamides C

2 ~ 8	1 ~ 4	-3 ~ 1

Region 20

(+)-aS-alterporriol C

-4 ~ 2	1 ~ 4	-1 ~ 1

Dehydroxyaquayamycin **Region 22** Sydowiols A

BTZ043 VI9376 37790

PA824 DNB1

-4 ~ 2	-2 ~ 1	1 ~ 3

Region 16

Ophiobolin K AU1235

SQ109

2 ~ 8	1 ~ 4	-1 ~ 1

Region 23

8'-O-demethylisonigerone

2 ~ 8	1 ~ 4	1 ~ 3

Region 27

Brevianamide S

Beauvercin TMC207

2 ~ 8	-2 ~ 1	1 ~ 3

Region 17

Pitipeptolides A

-4 ~ 2	1 ~ 4	-3 ~ 1

Region 19

Isoniazid Ethionamide Thioacetazone

Aminosalicylic acid Prothionamide Pyrazinamide

-4 ~ 2	1 ~ 4	1 ~ 3

Clofazimine BM212 C215 **Region 25** OPC67683

Figure 2.10 (Continued)

Firstly, the injectable and nature-derived TB drugs comprised of streptomycin, kanamycin, amikacin, capreomycin and viomycin clustered together in region 2 and no representative mycobacteria-active natural products occupied this region of physicochemical space. It is important to note that all of these drugs are protein synthesis

inhibitors. Another region that lacked representation of mycobacteria-active natural products from marine microbes is region 19, which contains five cell wall inhibitors (isoniazid, pyrazinamide, thioacetazone, ethionamide and prothionamide) and aminosalicylic acid. The low density of natural products in these regions indicates a promising field for new natural products with similar mode of mechanism with current drugs. Secondly, there were eight regions that had only mycobacteria-active natural products, region 9 in particular, the second biggest region with nine natural products of which four exhibited anti-*M. tuberculosis* MIC < 1 μM. These active natural products have unique properties in chemical space, as well as excellent biological activities. Finally, seven regions contained TB drugs, as well as mycobacteria-active natural products, and according to the theory that compounds with similar activity profile and chemical properties often show a similar mode of action, it is possible to predict the putative mode of action of these natural products. For example, region 10 contained TB drugs with small sizes and cell wall inhibition activities, as well as eight marine microbe-derived mycobacteria-active natural products (Figure 2.12).
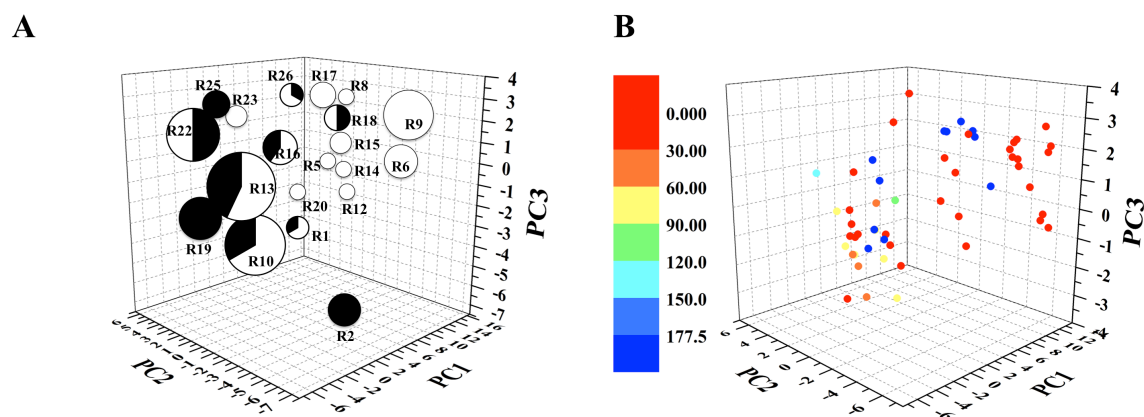
Figure 2.11 Score plot of mycobacteria-active natural products from marine microbes and TB drugs and candidates in regions. A. Total numbers of compounds in each region is indicated by sizes of pie charts, percentages of TB-active NPs from marine microbes and TB drugs in each region are showed in white and black, respectively. B. Anti-TB activity of 60 marine microbial NPs, color coded according to the legend (blue lowest and red highest activity). PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity).
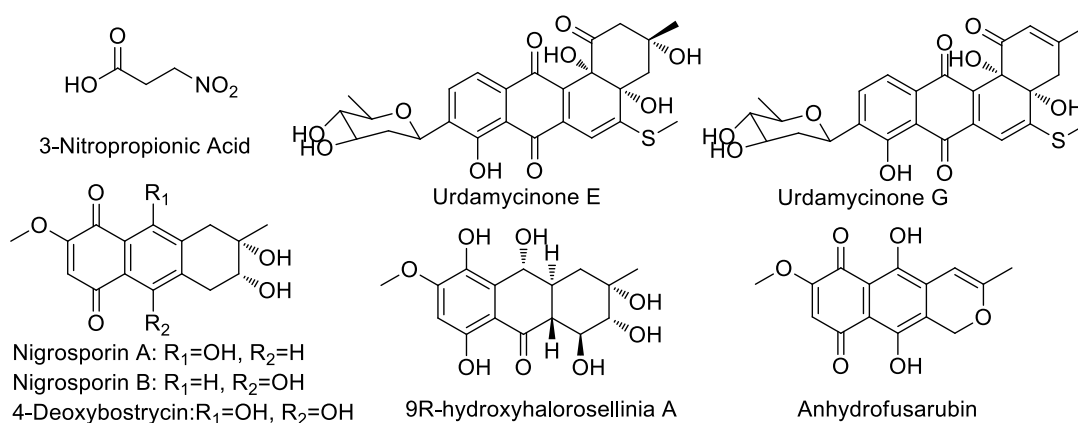
Figure 2.12 Structures of eight TB-active marine microbial natural products in region 10.

These natural products may be potential mycobacteria cell wall inhibitors. Similarly, natural products in region 22 and 26 could be predicted as promising cell wall inhibitors and ATP synthesis inhibitors, respectively. Moreover, there were twelve mycobacteria-active natural products and nine anti-TB drugs in region 13, including six DNA inhibitors and three protein inhibitors. Based on the same hypothesis, the mode of mechanisms of these compounds would be predicted as inhibitors of DNA or proteins of mycobacteria. This hypothesis was strongly supported by the fact that nine compounds in region 13 with known mode of mechanisms are all protein inhibitors, consistent with our prediction.

2.3.5 TB-active natural products isolated from endophytes

The second set of compounds was also extracted from the literature and has been reported to exhibit anti-TB activity. Compared with marine compounds, TB-active compounds from endophytes are relatively smaller and more aromatic. Then the compounds were analyzed by different regions according to their different values of PC1, PC2 and PC3 and Figure 2.13 showed some examples in each region.

Regions 1, 4 and 7 contained compounds with a combination of smaller size in the negative direction of PC1 and less aromatic features in the negative direction of PC2, but with the largely different lipophilic properties. Regions 19, 20 and 21 include compounds that are rather aromatic and polar, on the contrary, compounds in regions 7, 8 and 9 are characterized by higher polarity with decreased aromatic rings. Figure 2.14B showed the 4D plot of the 66 compounds by adding their anti-TB activity into consideration.

Similar with that of marine microbes, there are several regions that contained only clinic drugs but no TB-active NPs produced by endophytes, such as region 3 and region 25,
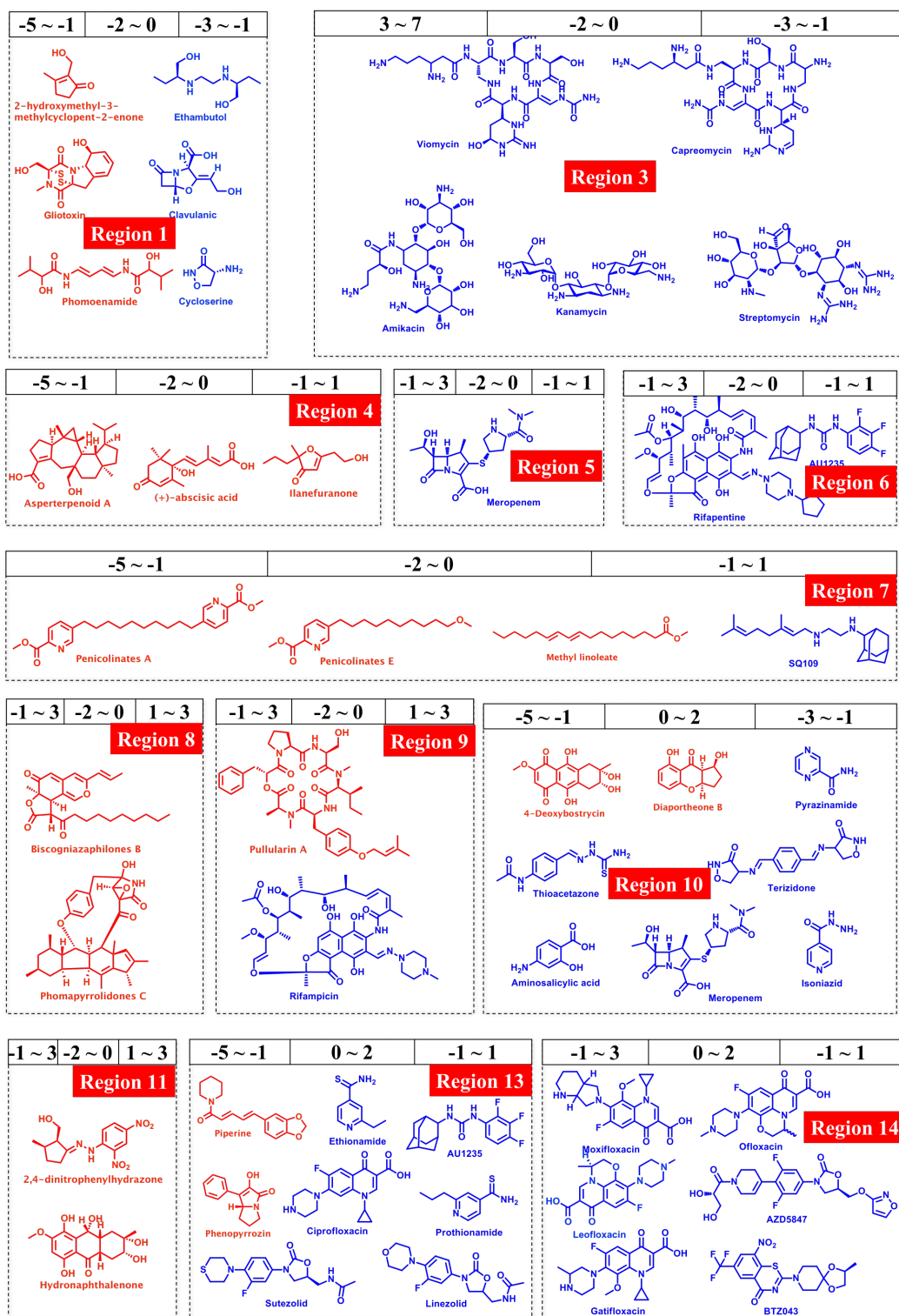
Figure 2.13 Score specifiction of the endophytes regions with typical structure examples. TB drugs and candidates are in blue and active marine microbial NPs are in red, PC1, PC2 and PC3 are shown on the top of each region.
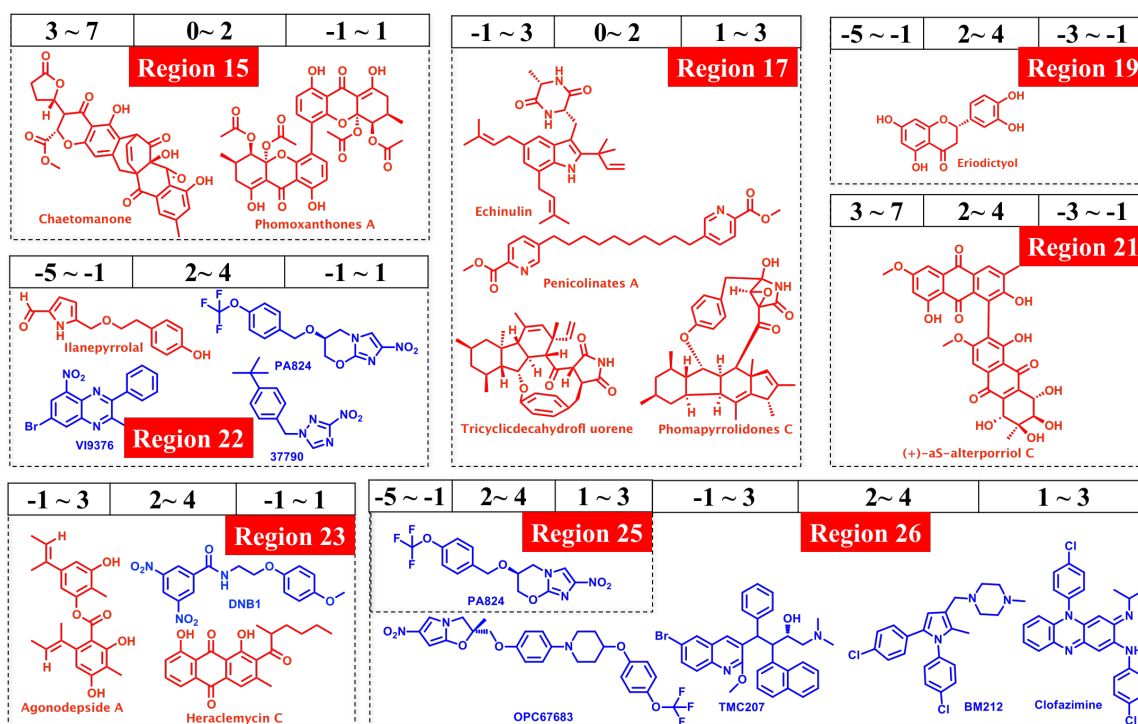
Figure 2.13 (Continued)
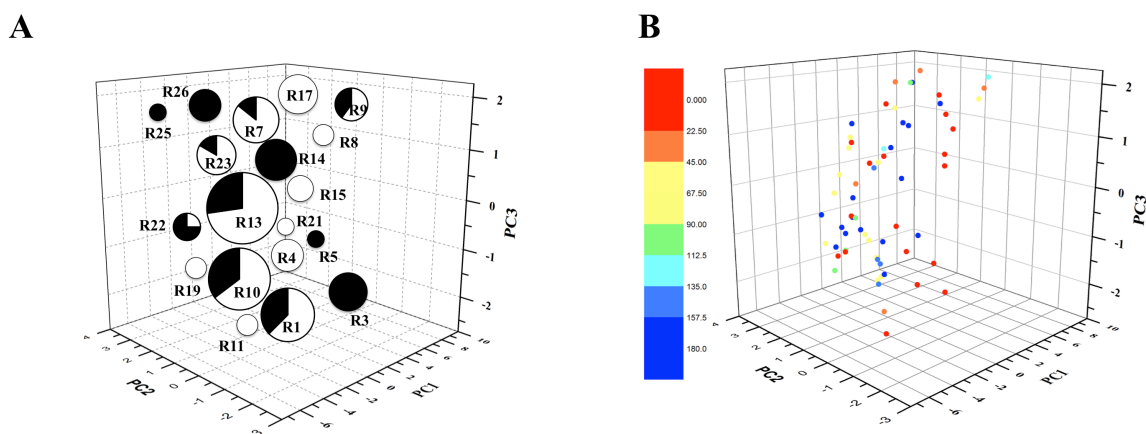
Figure 2.14 Score plot of mycobacteria-active natural products from endophytes and TB drugs and candidates in regions. A. Total numbers of compounds in each region is indicated by sizes of pie charts, percentages of TB-active NPs from endophytes and TB drugs in each region are showed in white and black, respectively. B. Anti-TB activity of 66 endophytes NPs, color coded according to the legend (blue lowest and red highest activity). PC1 (molecular size) versus PC2 (molecular romaticity) versus PC3 (molecular lipophilicity).

both with extreme PC values (Figure 2.14). Comparing with TB-active marine NPs and TB drugs, most of endophyte compounds are more "moderate" in ChemGPS space. Also, the endophytes NPs in region 9 and 10 could be predicted as possible RNA synthesis and

cell wall inhibitors, respectively, based on the mode of mechanisms of TB drugs in the same region. Particularly, the three most active compounds, with MIC values of 1.2 nM, 0.06 μM and 0.14 μM, respectively, were found to be quite small and polar. Two of them were located in region 1, which also contained three TB drugs with cell wall inhibition ability. Besides region 1, two compounds in region 13 that are non-polar and have no aromatic rings also exhibited significant activity. The least active regions in endophyte compounds were region 7 and region 10, which contained small and polar molecules that similar to region 1 but with increasing aromaticity.

2.4 Natural products near neighbors of approved drugs

Calculation of Euclidean distances based on ChemGPS scores was found to be a useful tool to interpret results and for finding interesting NP inspired leads for drug discovery.[110] Euclidean distances (EDs) of TB approved drugs and TB-active natural products were calculated over eight dimensions. The EDs were calculated between points P = (p1, p2, ... , pn) and Q= (q1, q2, ... , qn) in the Euclidean n-dimensional space, as defined by the following expression:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$
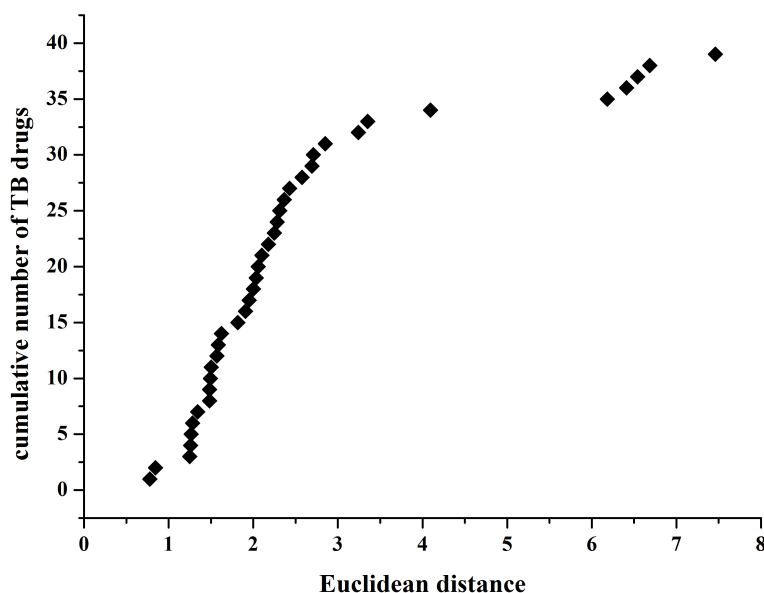
Figure 2.15 Distribution of ED to the nearest NP neighbor for the TB drugs and candidates. The cumulative number of drugs is plotted against the ED to the closest NP neighbor.

Thereby all NPs were assigned with 39 EDs, one ED to each drug. The NP/drug pairs were subsequently sorted in order of increasing EDs. In Figure 2.15, the 39 drugs are plotted against the ED to their closest NP neighbor. Interestingly, all drugs have a NP neighbor closer than ED = 8, and 43.6% of the drugs have a NP neighbor closer than ED = 2. This forms a strong argument that NPs from marine microbes and endophytes have the potential to serve as an important source of TB drugs. The structures of the top 10 shortest EDs of TB-active NPs/drugs were given in Figure 2.16.
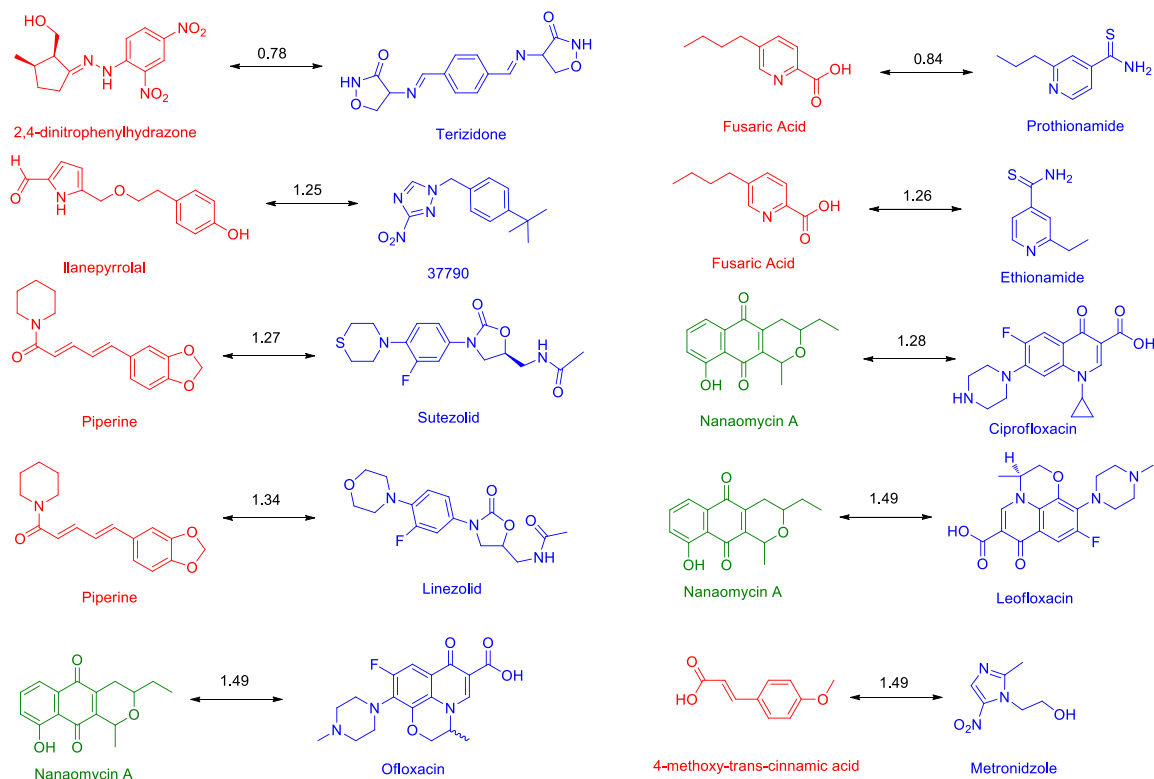
Figure 2.16 Structures of top10 shortest EDs of TB-active NPs/drugs pairs. Natural products from endophytes are in red, natural products from marine microbes are in green and TB drugs and candidates are in blue and EDs are given in black numbers.

Some natural products show short EDs with more than one drug neighbor (Figure 2.17). 2,4-dinitrophenylhydrazone, produced by the endophytic fungus mitosporic *Dothideomycete* sp. LRUB20,[123] is the nearest NP neighbor of four drugs, namely terizidone (ED = 0.78), thioacetazone (ED=2.00), DNB1 (ED = 2.10) and AZD5847 (ED = 2.43). Besides 2,4-dinitrophenylhydrazone, asterric acid is also a close neighbor of AZD5847 (ED = 2.70) and DNB1 (ED = 2.27). DNB1 also had another close NP neighbor *N*-trans-feruloy-3-*O*-methyldopamine, isolated from the endophytic fungus
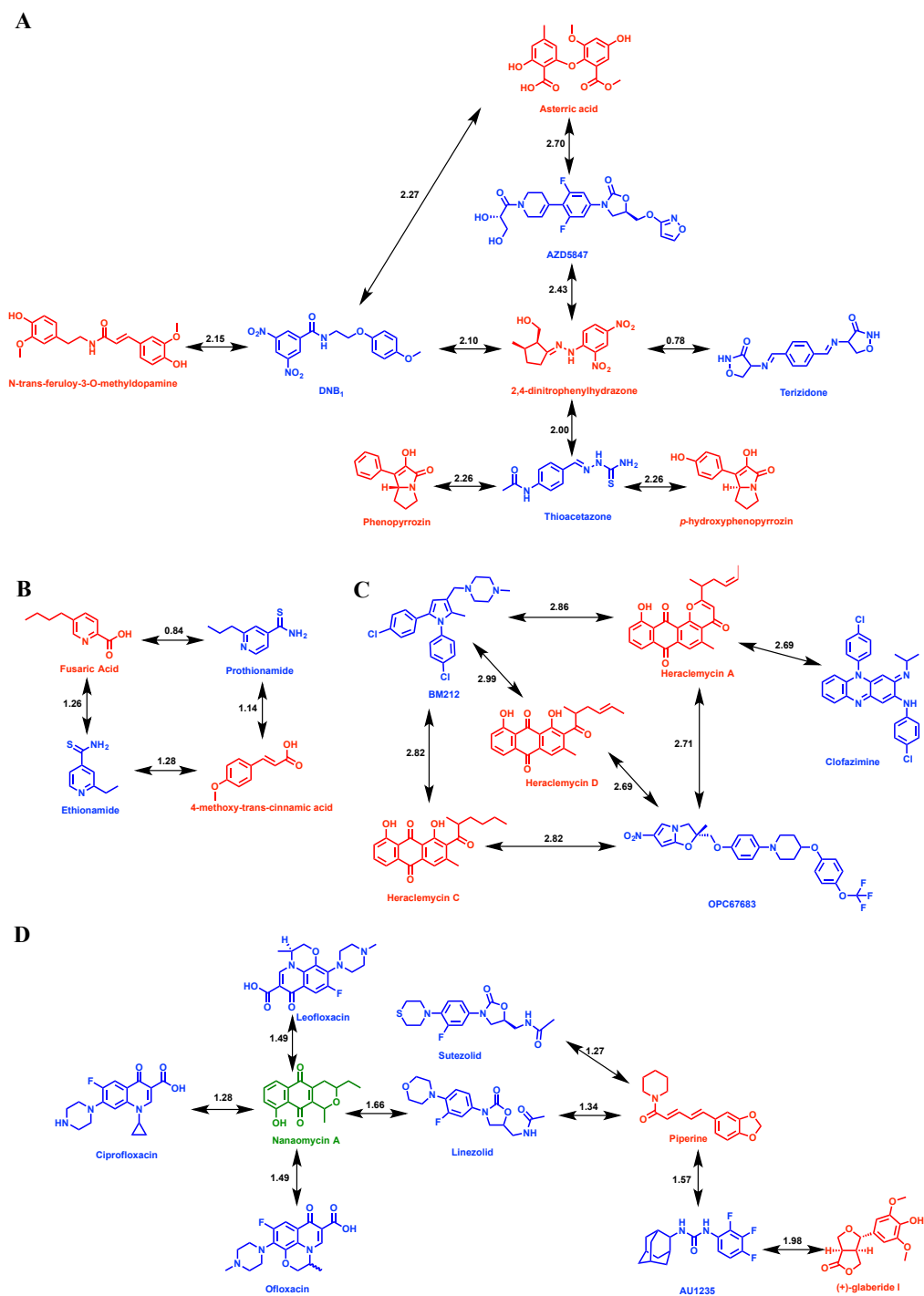
Figure 2.17 Structures of TB-active NPs/drugs pairs. Natural products from endophytes are in red and TB drugs and candidates are in blue and EDs are given in black numbers. A. 2,4-dinitrophenyl-hydrazone and its drug neighbors, B. fusaric acid and 4-methoxy-trans-cinnamic acid and their drug neighbors, C. heraclemycins and their drug neighbors and D. piperine and glaberide I and their drug neighbors.

Biscogniauxia formosana and exhibiting anti-mycobacterial activity against *Mycobacterium tuberculosis* strain H37Rv, with MIC value of 12.5 μg/mL.[77] Also, thioacetazone had two close NP neighbors: phenopyrrozin and *p*-hydroxyphenopyrrozin that showed very similar structures. Structures of these drugs and their close NP neighbors are given in Figure 2.17A.

Another example of an interesting drug/NP pair captured by this method is fusaric acid, isolated from the endophyte *Fusarium* sp.,[124] which has an ED of 0.84 to prothionamide and 1.26 to ethionamide. Also, compound 4-methoxy-trans-cinnamic acid from *Biscogniauxia formosana*[77] showed short EDs to those two drugs, 1.14 and 1.28 respectively. The structures are depicted in Figure 2.17B.

The novel anti-TB natural products from endophyte heraclemycin A, C and D showed close relationship with OPC67683 and BM212, with ED value between 2.5-3.0. Interestingly, OPC67683 and BM212 were found to be effective cell wall inhibitors, which could suggest a possible mode of action of the heraclemycins. Structures of these drugs and their close NP neighbors are given in Figure 2.17C.

The cell wall inhibitor TB drug candidate AU1235 has two active NP neighbors: (+)-glaberide I derived from endophyte *Annulohypoxylon ilanense* with anti-TB activity of 52 μg/mL (ED = 1.98);[125] piperine isolated from *Periconia* sp. with anti-TB activity of 1.74 μg/mL (ED = 1.57).[126] Piperine is also a close neighbor of two protein synthesis inhibiting drug candidates linezolid (ED = 1.34) and sutezolid (ED = 1.27). Structures of these drugs and their close NP neighbors are given in Figure 2.17D.

Nanomycin βA and αA, isolated from a marine derived *Streptomycetes* sp., have been reported to inhibit mycoplasma, fungi and Gram-positive bacteria, showing inhibitory activity against *M. tuberculosis* H37Rv with an MIC value of 8.0 μg/mL.[127] Nanomycins showed a close relationship with nine anti-TB drugs, including one cell wall inhibitor, five DNA inhibitors and two protein synthesis inhibitors, with ED value < 2 (Figure 2.18). We suggest that the natural products with short EDs to known drugs should be paid particular attention in developing new anti-TB drugs candidates because of their high similarity with known drugs in physico-chemial space.
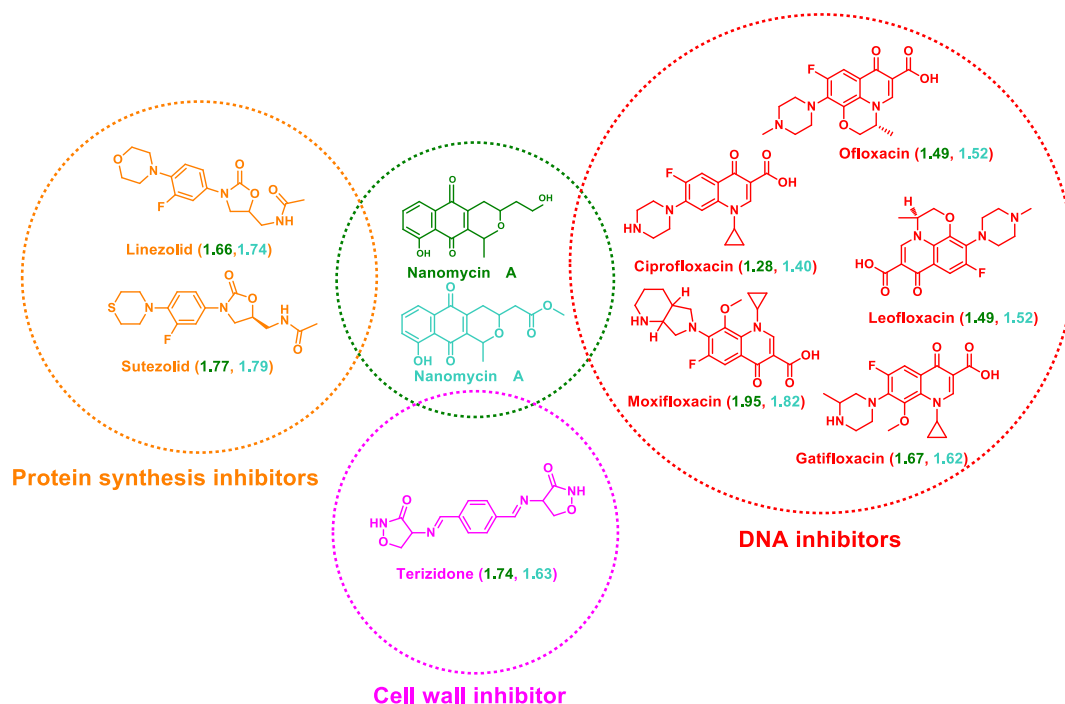
Figure 2.18 Structures of nanaomycin αA and nanaomycin βA and their close drug neighbors. Anti-TB drugs with protein synthesis inhibitor activity are in orange, DNA inhibitors are in red and cell wall inhibitor is in magenta. EDs between drugs and nanaomycin αA and nanaomycin βA are given in blue and cyan numbers, respectively.

2.5 Natural products from an in-house library of TB-active compounds

The fourth dataset consists of 157 molecules, which came from the natural products library (contains more than 1440 compounds) of Prof. Lixin Zhang's lab. These compounds were tested in a HTS BCG assay, evaluating 48 active and 109 inactive compounds. They were identified from different organisms, including marine microbes, endophytes, land strains and plants, where they are represented as green, red, blue and orange spheres, respectively (Figure 2.19). Because of their varied structures, the compounds from marine strains span a comparably large sector of the chemical property space, in particular along PC1 and PC2. The plant compounds are mapped further out in the negative direction of PC1 and PC2, indicating smaller size and higher polarizability. Interestingly, in this study, compounds from endophytes clustered together with the compounds from plants. This confirms a long-standing suggestion that the endophyte compounds show a "chemical similarity" to the plant compounds by a complex system of communication. The natural products from land strains show similar properties with the

marine compounds in term of PC2 and PC3 but are relative small (negative PC1).
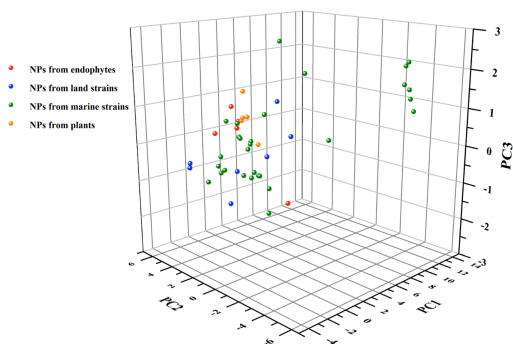
A   **B**
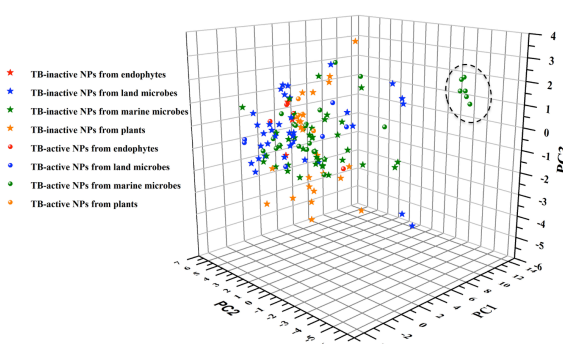
Figure 2.19 Score plot of natural products from an in-house library. A. PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 5 natural products from endophytes (red), 7 natural products from land strains (blue), 30 natural products from marine strains (green) and 6 natural products from plants (orange) and B. PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 48 TB-active natural products (sphere) and 109 TB-inactive natural products (star), 6 unique TB-active compounds from marine microbes are circled.

Actinomycin $X_{0\beta}$: R=OH
Actinomycin X_2: R=O
Actinomycin D: R=H

Lobophorin A: R_1=NH$_2$, R_2=OH
Lobophorin B: R_1=NO$_2$, R_2=OH
Lobophorin G: R_1=NH$_2$, R_2=OAc
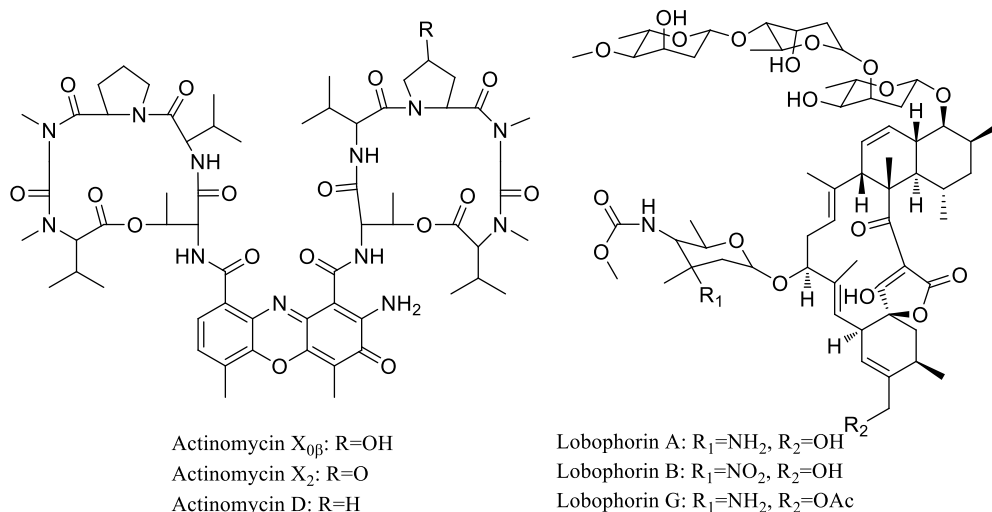
Figure 2.20 Structures of six TB-active compounds from Zhang's lab with extremely different PCs.

In general, the physico-chemial properties of TB-inactive and TB-active compounds tested in this bioassay are largely overlapping, i.e. compounds from endophytes cluster together with each other in Figure 2.19B. However, there are still some significant differences between the active and inactive ones. The inactive compounds apparently occupy a much larger area along PC3 than the active compounds, indicating that appropriate degree of polarity may be essential for the anti-TB activity. Furthermore, six

active compounds from the marine dataset (circled in Figure 2.19B) showed extremely different PCs values with any of the inactive compounds. This compound group is set aside by a much higher prediction score in PC1 and lower values in PC2, primarily depending on size and aromatic parameters. In addition to the aberrant physico-chemial properties demonstrated for this compound group, it is also concluded from the bioassay data that it is one of the most active compound group within the active compounds identified. Structures of these compounds are presented in Figure 2.20.

2.6 Prediction of the anti-TB activity among marine compounds
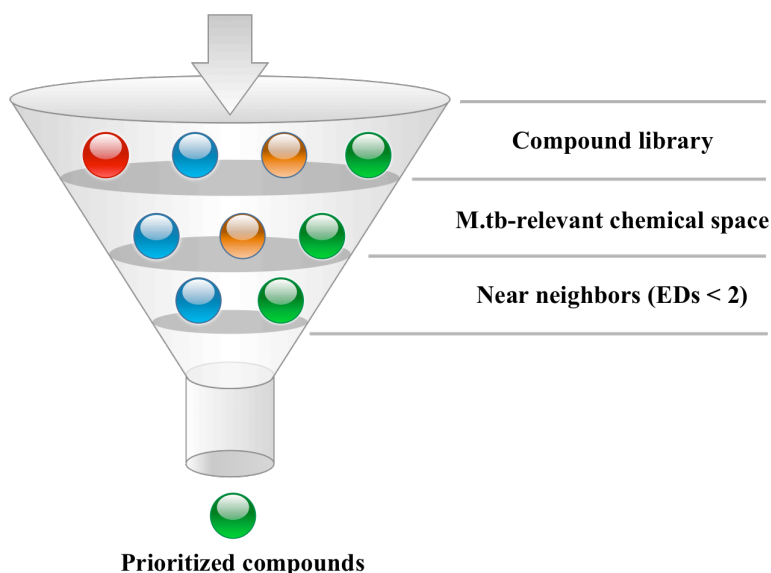
Figure 2.21 Main factors to design diverse combinatorial libraries or to select diverse compounds to conduct anti-*M.tb* screening collection.

The above information can be used to develop a strategy to select compounds when conducting anti-TB screening (Figure 2.21). First, mapping of candidate compounds into the TB-relevant chemical space defined by TB drugs and TB-active natural products allows for a quick and efficient filter for potential TB-activity. The location of compounds in this model can also be used to predict potential mode of action (MoA). Second, compounds with EDs < 2 to known TB drugs provides a second filter. Screening of the prioritized compounds in chemical space may directly lead to the identification of biologically relevant structures for TB drug discovery.

Data set 5 (Figure 2.2) was analyzed using the model. 327 of 336 compounds were found to fall into the 27-region chemical space (Figure 2.22). Similar to TB-active natural products, region 13 contained the most untested compounds (90). Surprisingly, three

regions 4, 11 and 24, which have no represented TB drugs nor TB-active compounds, occupied with 5, 1 and 2 untested compounds, respectively. These compounds may show unique structure properties.
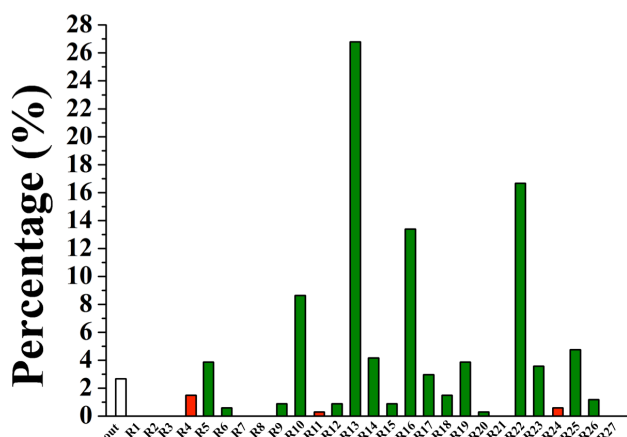
Figure 2.22 Histogram of the percentage of untested NPs from marine microbes in each regions. Compounds not falling into the 27 regions were plotted in black, compounds in regions 4, 11 and 24 without represented TB drugs and TB-active NPs were plotted in gray.
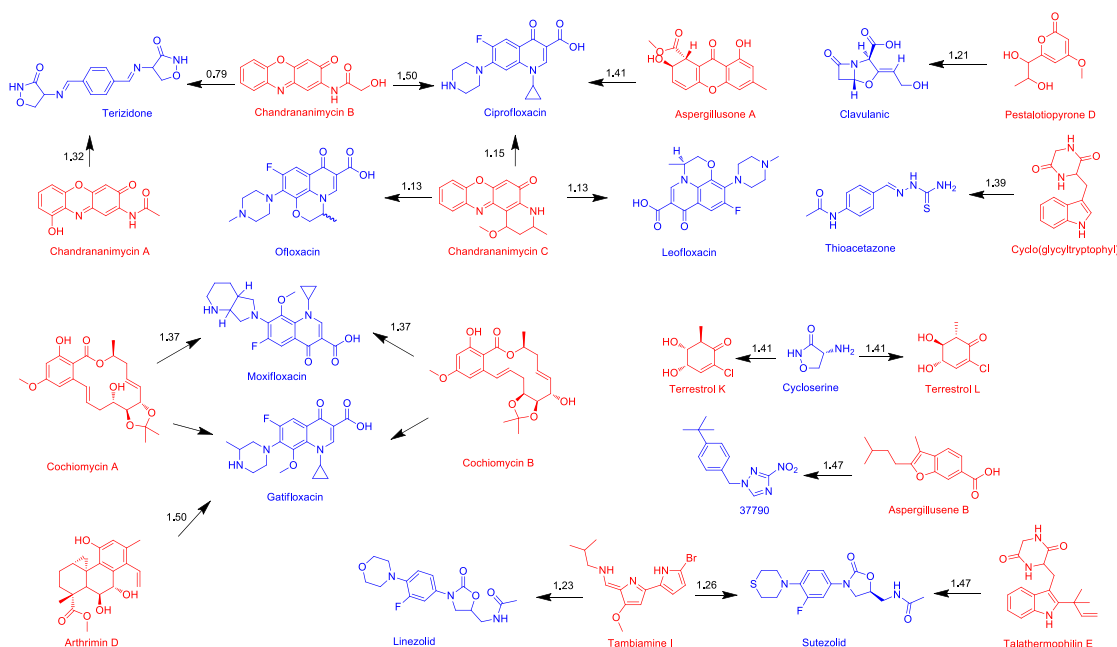
Figure 2.23 Structures of top 20 shortest EDs of untested natural products/drugs pairs. Untested natural products from marine microbes are in red and TB drugs and candidates are in blue and EDs are given in black numbers.

Filtering these compounds using EDs calculation to identify near neighbors with known TB dugs produced 82 prioritized structures with ED < 2 to at least one drug. Structures of

the top 20 nearest EDs of untested natural products/drugs are given in Figure 2.23. Untested compounds showing short EDs with more than one drug neighbor may be worth screening.

2.7 Conclusion

Tuberculosis threatens people's lives around the world and the appearance of drug resistance cases has increased the need to identify novel anti-TB drugs. With the aim to demonstrate the enormous chemical diversity of nature, especially that of the marine microbes and endophytes, we used Instant J Chem and the chemical space navigation tool ChemGPS-NP to compare sets of marine microbe-derived compounds, endophyte compounds and TB drugs with respect to physico-chemial properties and their occupation of the chemical space.

We analyzed the physico-chemial properties with respect to the Ro5 and tPSA parameters of 39 TB drugs and candidates, 60 mycobacteria-active natural products, and 336 untested natural products. While Ro5 gives some information about cell permeability, the complex etiology of the disease technology involving, oral bioavailability, lung-alveolar macrophages, and mycobacterial cell permeability, means these predictions are not yet sophisticated enough to be taken with any degree of certainty. Analysis showed a wide area of acceptable physico-chemial properties with log P -1 to +5, probably due to the need to address oral bioavailability in conjunction with penetration through the unique cell wall of *M.tb*. Further analysis by ChemGPS-NP defined 27 regions within TB active space. Based on the theory that compounds with similar activity profile and chemical properties often show similar mode of action, location of a new compound with a TB drug may indicate a potential active compound. Moreover, EDs to known TB drugs may be a better predictive tool. We concluded that if a compound has a ED of <5 to any of the current TB drugs, then the compound has a much higher chance of itself being active.

Chapter III Whole cell-based screening of natural products and microbial extracts for anti-TB activity

Abstract: This chapter presents a strategy for using Mycobacteria bovis BCG as model to detect anti-TB activities in pure natural products and natural products extracts. Antimicrobial screening of a small collection of 140 natural products resulted in the identification of promising anti-infectious drug leads. Additionally, a total of 2562 microbial crude extracts were generated from 654 actinomycetes and then screened against BCG. Furthermore, 415 positive hits have been evaluated by HPLC.

3.1 Introduction

Diversifying the microbial sources of natural compounds is of primary concern in library development. Microorganisms of both terrestrial and marine origin have proven to be excellent sources of novel natural products, and recent advances in microbial genomics (genome sequencing, microbial ecology and metagenomics) have unequivocally demonstrated that the biosynthetic potential of natural products in bacteria is much higher than previously appreciated.[128] It was reported that the bacterial species of *Micromonosporineae*, *Pseudonocardineae*, *Streptomycineae* and *Streptosporangineae* are the most frequents sources of novel secondary metabolites.[129] Moreover, species inhabiting unique environments, such as marine microbes and endophytes, have been proposed to be promising resources of novel compounds.[130]

IMCAS natural product library (NPL) is a collection of over 35,000 samples from 12,000 microorganisms, including over 4,000 actinomycetes isolated from the sediment from the South China Sea down below 4,000 m and more than 14,000 marine crude extracts for screening against infectious diseases.[131] Value adding to the biota samples was achieved via systematic alteration of the culture media. In this project, 2562 crude extracts were generated from 654 actinomycetes, including 329 land strains, 200 endophytes and 125 marine microbes collected from various extreme environments of China grown under different media (Figure 3.1).
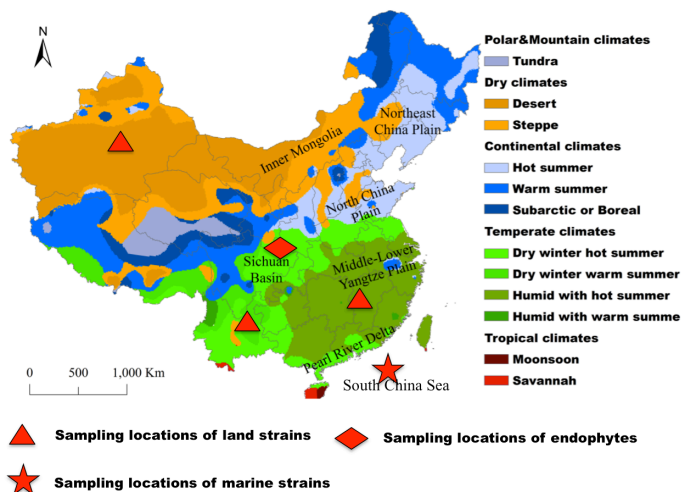
Figure 3.1 Sampling locations of 329 land strains, 200 endophytes and 125 marine microbes.

A serious shortcoming of enzyme-based screening assays is that while natural products, such as fosmidomycin, may show good inhibition on pure enzymes, they showed weak

inhibitory effect on the growth of *M. tuberculosis* due to their inability to penetrate through the membrane.[132] To screen the large numbers of crude extracts, fractions and compounds for growth inhibition, simple screening techniques are needed. A whole cell-based screening assay was developed to screen the effect of microbial crude extracts, fractions as well as pure compounds on the growth of *Mycobacterium bovis* ATCC35743 (BCG) (a slow-growing non-virulent strain closely related to *M. tuberculosis* H37Rv in terms of drug susceptibility profile and genetic composition).[133]

3.2 Methodology

3.2.1 Whole cell-based screening assay

The BCG strain was transformed with green fluorescent protein (GFP) constitutive expression plasmid pUV3583c with direct readout of fluorescence as a measure of bacterial growth. BCG was grown at 37 °C to mid log phase in Middle brook 7H9 broth (Becton Dickinson) supplemented with 10% OADC enrichment (Becton Dickinson) 0.05% tween-80 and 0.2% glycerol, which was then adjusted to $OD_{600}= 0.025$ with culture medium as bacterial suspension. Aliquots (80 μL) of the bacterial suspension were added to each well of the 96-well plates (clear flat-bottom), followed by adding compounds or extracts (2 μL suspended in DMSO), which were serially diluted twofold. Isoniazid served as positive control and DMSO as negative control. The plates were incubated at 37 °C for 3 days, and GFP fluorescence was measured with Multi-label Plate Reader using the bottom read mode, with excitation at 485 nm and emission at 535 nm. MIC was defined as the minimum concentration of drug that inhibits more than 90% of bacterial growth reflected by fluorescence value (Figure 3.2).
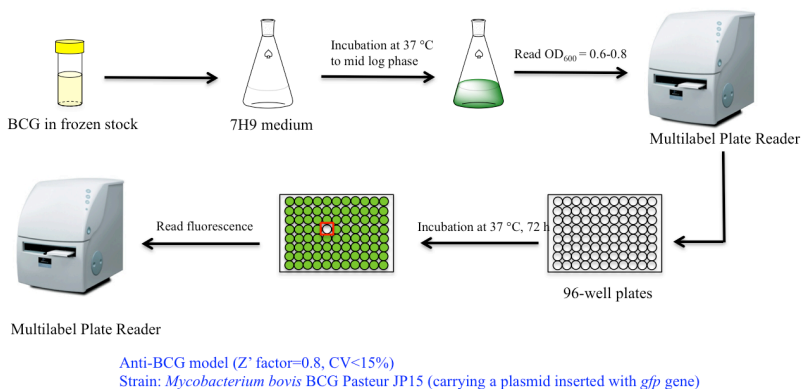
Figure 3.2 SOP of whole cell-based anti-BCG screening.

3.2.2 Construction of diversified natural products library from microbes

Natural product library was constructed based on the protocol, which includes cultivation, extraction, activity screening and HPLC analysis (Figure 3.3). All of the land strains and marine strains were cultured in 3 different media and all of the endophytes were cultured in 6 media (Table 3.1). With the purpose of selecting potential strains with diverse chemical constituents and strong biological activities, HPLC analysis and anti-microbial screening were conducted for 2562 generated crude extracts.
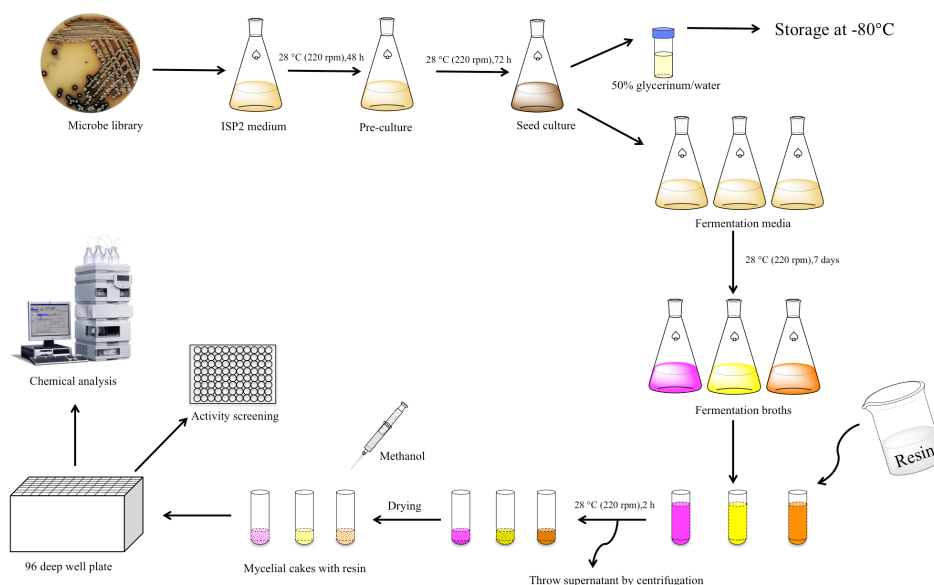
Figure 3.3 SOP for construction of the microbial natural products library.

Table 3.1 Different fermentation media used in microbial crude extracts library

Media	329 land strains	125 marine strains	200 endophytes
AM2	√	√	√
NM2	√	√	√
MPG	√	√	√
M001	X	X	√
M12	X	X	√
M21	X	X	√

3.2.3 HPLC evaluation of active crude extracts

Around 20 mg of each samples was dissolved in 500 μL methanol and 20 μL was injected

to reversed phase HPLC. HPLC separations were performed on a Zorbax SB-C$_{18}$ column (4.6 mm × 150 mm) using conditions that consisted of a linear gradient from 95% H$_2$O /5% MeCN to 100% MeCN in 20 min at a flow rate of 1 mL/min; held at MeCN for a further 5.0 min; then a linear gradient back to 95% H$_2$O/5% MeCN in 0.1 min, then held at 95% H$_2$O/5% MeCN for 4.9 min. The total run time for each extract injection was 30 min.

3.3 Result

3.3.1 Screening of pure compounds

Apart from crude extracts, 140 natural products from colleagues and collaborators were also evaluated for their activities against infectious bacteria, including *M. bovis* BCG, *Staphylococcus aureus* (SA), methicillin resistant *Staphylococcus aureus* (MRSA), *Bacillus subtilis* (BS) and *Pseudomonas aeruginosa* (PA). The compounds comprise about 10 different classes of small molecules that reflect a substantial chemical diversity even in a relatively small collection. A pie chart, showing the distribution of the 140 compounds is shown in Figure 3.4. A majority of the compound samples are terpenoids (diterpenoids and sesquiterpenoids), constituting 55% of compound samples. This is followed by polypeptides (10%), steroids (7.86%), alkaloids (7.14%), phenylpropanoids (5.71%), lignans (3.57%), quinones (3.57%), bisnaphthalenes (2.14%), flavonoids (1.43%), other kinds of compounds (1.6%), tetracyclics (1.43%) and polyamides (0.71%).
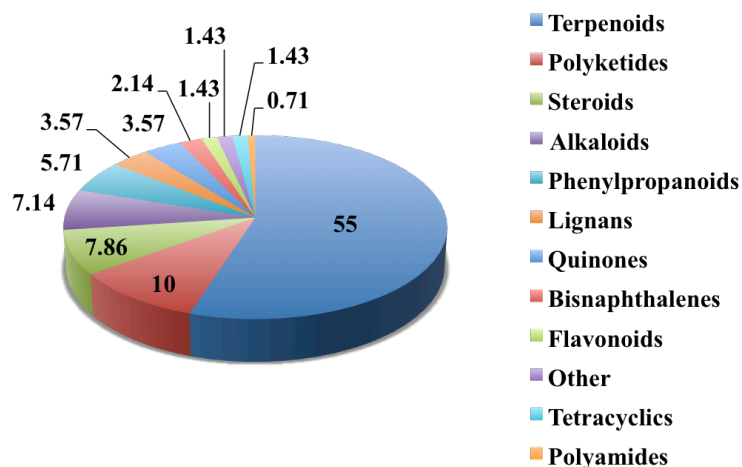
Figure 3.4 Distributions of 140 natural products by compound type.

Among 140 screening compounds, large proportions of tested compounds showed anti-

A

B
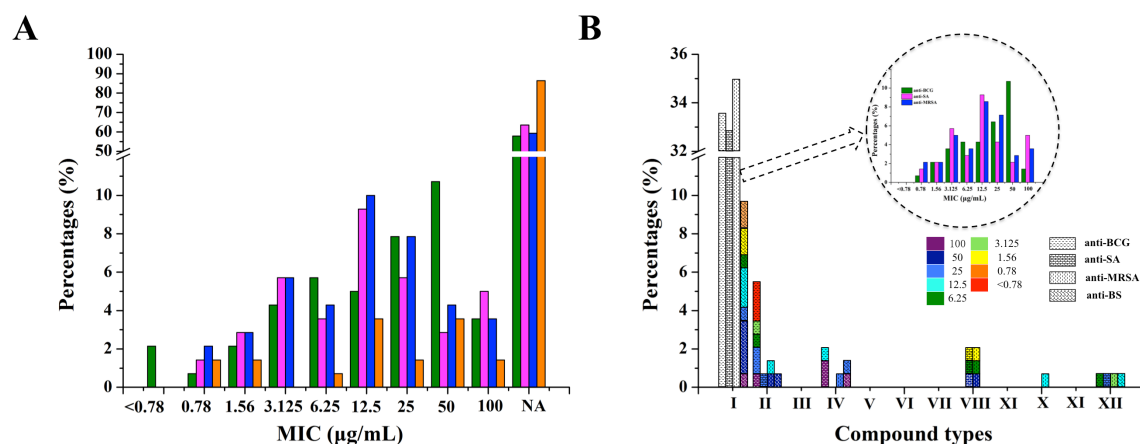
Figure 3.5 Histograms of the percentage of anti-microbial activities of 140 natural products. A. Activity results are shown by MIC values, anti-BCG activity (olive), anti-SA activity (magenta), anti-MRSA activity (blue) and anti-BS activity (orange). NA: not active. B. Activity results are shown by different compounds types, type I: terpenoids, type II: polyketides, type III: steroids, type IV: alkaloids, type V: phenylpropanoids, type VI: lignans, type VII: quniones, type VIII: bisnaphthalenes, type IX: flavonoids, type X: others, type XI: tetracyclics, type XII: polyamides. MICs are shown in different colors.

Caesanine A
MIC_{BCG}=50 $f{\subset}g$/mL
MIC_{SA}=12.5 $f{\subset}g$/mL
MIC_{MRSA}=12.5 $f{\subset}g$/mL
MIC_{BS}=50 $f{\subset}g$/mL

Proximicin B
MIC_{BCG}=6.25 $f{\subset}g$/mL
MIC_{MRSA}=3.125 $f{\subset}g$/mL
MIC_{BS}=12.5 $f{\subset}g$/mL

Abyssomicin J
MIC_{BCG}=3.125 $f{\subset}g$/mL
MIC_{MRSA}=12.5 $f{\subset}g$/mL
MIC_{BS}=50 $f{\subset}g$/mL

Tanshinoside A
MIC_{BCG}=50 $f{\subset}g$/mL
MIC_{SA}=0.78 $f{\subset}g$/mL
MIC_{MRSA}=1.56 $f{\subset}g$/mL
MIC_{BS}=1.56 $f{\subset}g$/mL
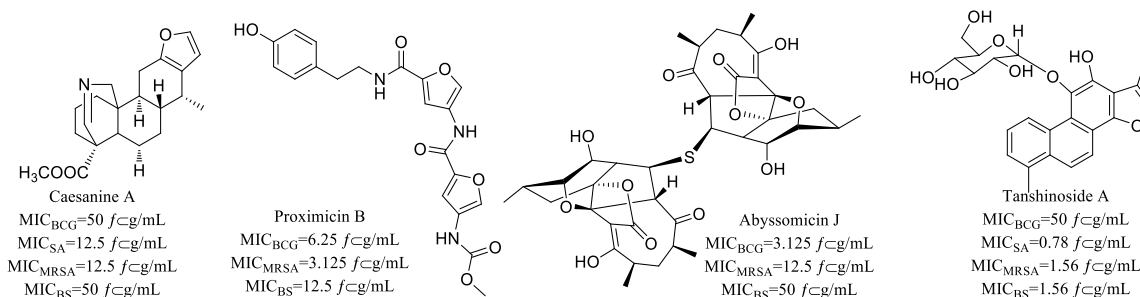
Figure 3.6 Structures of 4 example compounds.

BCG (42.14%), anti-SA (36.43%) and anti-MRSA (40.71%) activities, while only 13.57% of compounds had anti- BS activity (Figure 3.5A). However, none of them showed inhibition effects on Gram-negative pathogen *Pseudomonas aeruginosa*. Within this small collection of compounds, 3% of them displayed extremely strong anti-BCG activity with MIC values of 0.78 *μ*g/mL or even lower. The most potent compounds exhibited activity against SA and MRSA with MIC values of 0.78 *μ*g/mL. Among 12 compounds types, terpenoids contained the largest number of active compounds, while a relatively small part of them exhibited low MICs against tested models regarding to the high percentage it occupied in the whole test collection (Figure 3.5B). The most active anti-BCG compounds are found to be all polyketides isolated from a marine

actinomycetete, suggesting promising prospects of this particular compound type and organism source. Structures of 4 example compounds in the antimicrobial screening assays were shown in Figure 3.6.

3.3.2 Screening of microbial extracts

With the aim of identifying new anti-TB active microbial natural products, 2562 crude extracts generated from 654 actinomycetes were screened against *M. bovis* BCG using the above whole cell-based model. Activity result of these crude extracts was evaluated and showed in heatmap (Figure 3.7).
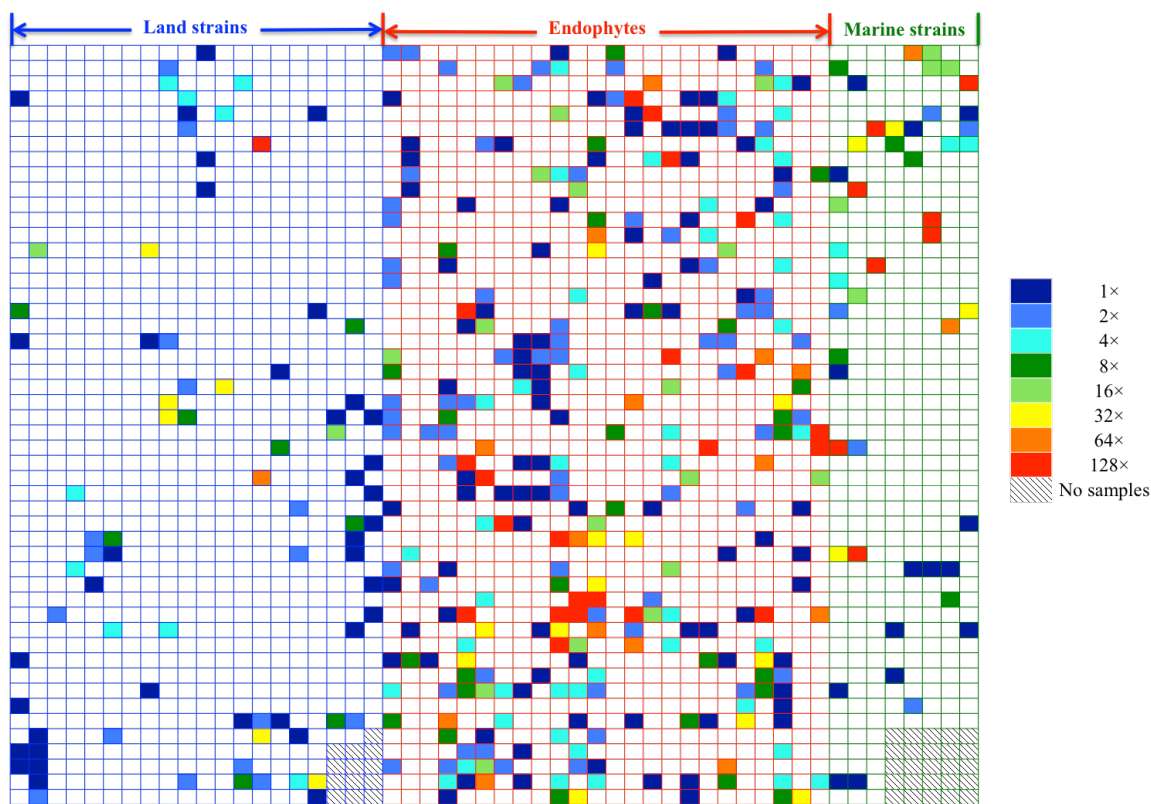
Figure 3.7 Heatmap of the anti-BCG screening on crude extracts. 2562 samples were prepared from 654 strains isolated from land strains (blue blanks), endophytes (red blanks) and marine microbes (green blanks), color coded according to the legend (blue lowest and red highest activity).

In total, 415 out of 2562 (16.2%) crude extracts showed positive results in anti-BCG screening assay (Figure 3.8). Around one third of the active extracts (33.73%) were effective on killing more than 90% BCG bacillus at their original concentration (100 μg/mL). With the increase of dilution factors, less extracts were able to play effective role in anti-BCG screening, only 5.3% and 3.86% of active microbial extracts with dilution

factors of 32 and 64. Almost 7% of active extracts were still active when they were diluted by 128 times, showing great potential for further chemical investigation. The high percentage of actives at 128-fold dilution may due to several strains showing strong anti-BCG activity while cultured under different media. For example, 6 crude extracts derived from the same endophyte ES120079 all exhibited activity with the dilution factor of 128.
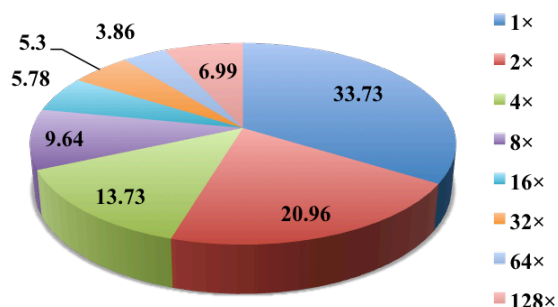
Figure 3.8 Pie chart of anti-BCG activities of 415 active crude extracts.

Percentages of active extracts from land strains, endophytes and marine strains were 8.1%, 23.9% and 12.8%, respectively. Overall, comparing with land strains, endophytes and marine actinomycetes were more likely to produce active metabolites. Almost half of active land strain extracts were only active against BCG at the used highest concentration and in total, only around 2.5% of them showed activity with dilution factors of 64 (1.25%) or 128 (1.25%) (Figure 3.9A). Even through a relatively large proportion of endophyte extracts exhibited anti-BCG activities, only around 20% of them showed activity with dilution factors larger than 8. Some reports proposed that within the evolutionary time, endophytes have adapted to their unique environments by genetic variation, including uptake of some plant DNA into their own genomes.[134, 135] Endophytes in this project were collected from various Traditional Chinese Medicine plants. This could have led to the ability of endophytes to biosynthesize phytochemicals that are characteristic of the respective host plants. It could also explain the high positive rates of endophyte extracts in anti-BCG screening assay and due to the moderate nature of TCMs in treating diseases, a relative small amount of endophytes extracts showed strong activity. Contrary to endophytes, approximate 15% of active marine microbial extracts displayed anti-BCG activity with dilution factors of 128, the second largest proportions among all active marine samples indicating the outstanding characteristics of marine actinomycetes as source for TB drug discovery.
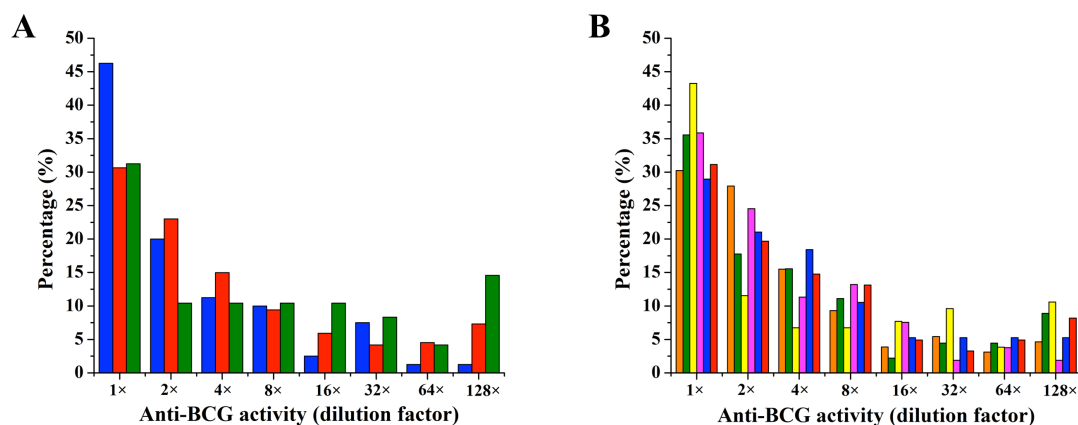
Figure 3.9 Histograms of the percentage of anti-BCG activities of 2562 microbial crude extracts. A. Biological sources: land strains (blue), endophytes (red) and marine strains (olive). B. Different media: AM2 (orange), NM2 (olive), MPG (yellow), M001 (magenta), M12 (blue) and M21 (red).

The characteristics of 6 culture media in anti-BCG metabolites production by 654 actinomycetes are shown in Figure 3.9B. Media M9, M001 and AM2 exhibited the most potential on inducing synthesis of active natural products with hit rates of 30.5 % (200 endophytes samples in total), 26.5 % (200 endophytes samples in total) and 19.7% (654 samples in total), respectively. Among all active samples, top 3 media with dilution factors larger than 8 were MPG, M21 and M12, with proportions of 31.7%, 21.3% and 21.1%, respectively. In our previous research, M21 was also reported as the most productive medium among 10 fermentation media.[131]
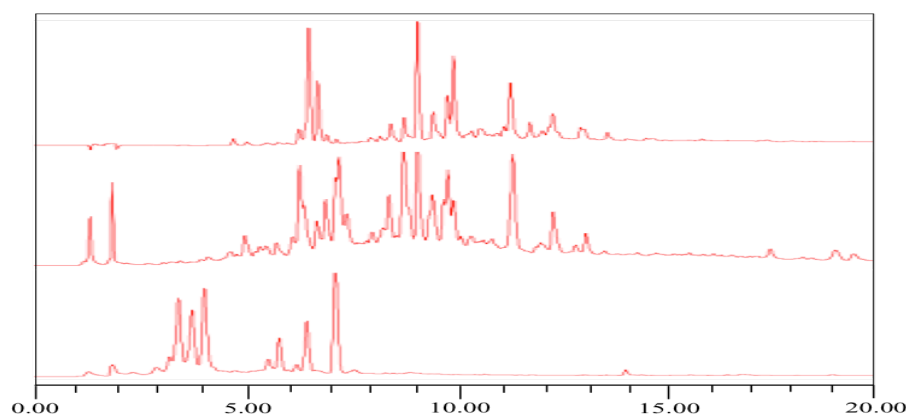
Figure 3.10 HPLC spectra of three active anti-BCG microbial crude extracts.

The 415 active microbial crude extracts were subsequently analyzed by HPLC. By considering the number of peaks as well the area of the peaks present in the chromatogram, promising samples were selected for further investigation (Figure 3.10).

3.4 Conclusion

In this study, we have described and implemented a cell-based HTS platform for discovering small molecules as well as crude extracts showing growth inhibition to *M. bovis* BCG. Using the validated assay, a small library consisted of 140 compounds was screened and 3% were identified as potential candidates that are capable of inhibiting BCG growth at very low concentrations (less than 0.78 μg/mL). With the aim to identify potent anti-TB natural products from actinomycetes, a microbial extract library was also generated from 654 strains and 2562 crude extracts were evaluated for their activity. Around 400 hits were identified from the library, which was the basis for the TB-focus work. Given consideration of anti-TB activity as well as chemical diversity, 51 microbial crude extracts derived from 37 actinomycetes were selected for further investigation, which include an endophyte (Chapter 4), 50 microbial crude extracts from endophytes, desert strains and marine strains (Chapter 5, 6 and 7).

Chapter IV Bioassay-guided isolation of an endophytic *Streptomyces* sp. Y3111 from Traditional Chinese Medicine

STATEMENT OF CONTRIBUTION TO CO-AUTHORED PUBLISHED PAPER

This chapter includes a co-authored paper. The bibliographic details of the co-authored paper, including all authors, are:

Miaomiao Liu, Wael M. Abdel-Mageed, Biao Ren, Wenni He, Pei Huang, Xiaolin Li, Krishna Bolla, Hui Guo, Caixia Chen, Fuhang Song, Huanqin Dai, Ronald J Quinn, Tanja Grkovic, Xiaoping Zhang, Xueting Liu*, Lixin Zhang*.

My contribution to the paper involved:

Cell-based screening of microbial crude extracts, fermentation and extraction, isolation and purification of compounds, structure elucidation, bioassay evaluation of pure compounds and manuscript writing.

(Signed) _____ (Date)_____
Miaomiao Liu

(Countersigned) _____ (Date)_____
Corresponding author of paper: Lixin Zhang / Xueting Liu

(Countersigned) _____ (Date)_____
Supervisor: Ronald J Quinn

Abstract: As part of a search for antitubercular substances from natural sources, we screened a library of endophytic microbes (50 strains and 300 crude extracts in total) isolated from traditional Chinese medicines (TCMs) for growth inhibitory activity against Bacillus Calmette-Guérin (BCG). The crude extract of Streptomyces sp. strain Y3111, which was associated with the stems of Heracleum souliei, showed good anti-BCG activity with an MIC value of 6.25 μg/mL. Bioassay-guided isolation led to four new pluramycin-type compounds, heraclemycins A–D (55 – 58). Their structures were determined by different spectroscopic techniques including HRESIMS, 1D NMR, and 2D NMR. This is the first report of pluramycin analogues produced by TCM endophytic microbes as well as the first example of BCG-selective pluramycins. Heraclemycin C (57) showed selective antitubercular activity against BCG with a MIC value of 6.25 μg/mL and a potential new mode of action.

4.1 Introduction

Natural products derived from microbes have been an important source of TB therapeutics. In order to minimize rediscovery of known compounds and access new and novel molecular scaffolds for drug discovery, rare and previously undiscovered microbial taxa sourced from extreme environments and specialized niches can be investigated. Plant endophytes represent one such source. Plants have developed highly diverse mechanisms to cope with their environment, and it is remarkable that all plants on earth harbor endophytes.[68] Many of them are capable of synthesizing bioactive compounds that can be used by the plant for defense against pathogenic fungi and bacteria. Endophytes represent a large variety of microbial adaptations that have developed in special and sequestered environments, and their diversity and specialized habitat make them an exciting field to search for new medicines.[136]

Thus, with the aim of identifying new anti-TB natural products from endophytes, a collection of 300 crude extracts from 50 endophytic microbes isolated from TCMs in Sichuan Province (Figure 4.1) was screened for growth inhibitory activity against *Mycobacterium bovis* BCG (chapter 3). 65 out of 300 (21.7%) crude extracts were identified as hits showing anti-BCG activity at the concentration of 100 μg/mL or lower (Figure 4.2).

Figure 4.1 Sampling information of 50 endophytes. A. The location of Sichuan Province in China; B. The sampling locations of studies endophytic microbes from TCMs on Sichuan plateau, a. Yajiageng, Kangding County, b Laoyulin, Kangding county, c Zheduo Mountain, Kangding County, d Bajiaolou, Yajiang County, and e Longdeng, Daofu County.
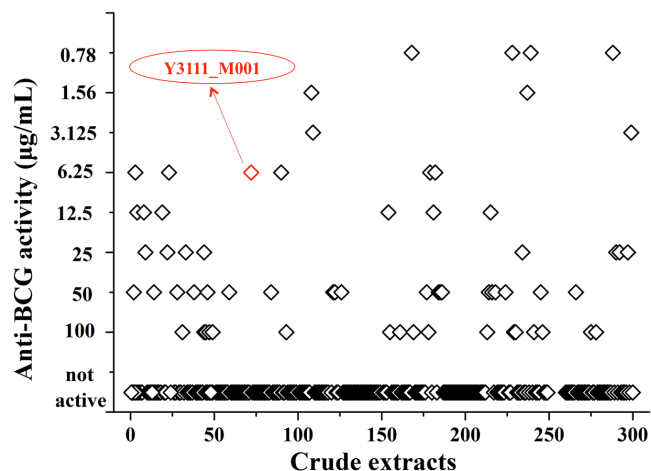
Figure 4.2 Result of the anti-BCG screening on 300 crude extracts from 50 endophytes.

Chemical evaluation for the top 5% (15) active crude extracts was conducted by comparison of the HPLC spectra (Figure 4.3). The crude extract derived from an endophyte Y3111 in medium M001 (MIC value of 6.25 μg/mL) displaying the most various chemical diversity under HPLC analysis, thus, was selected for large scale fermentation and identification of the constituents.
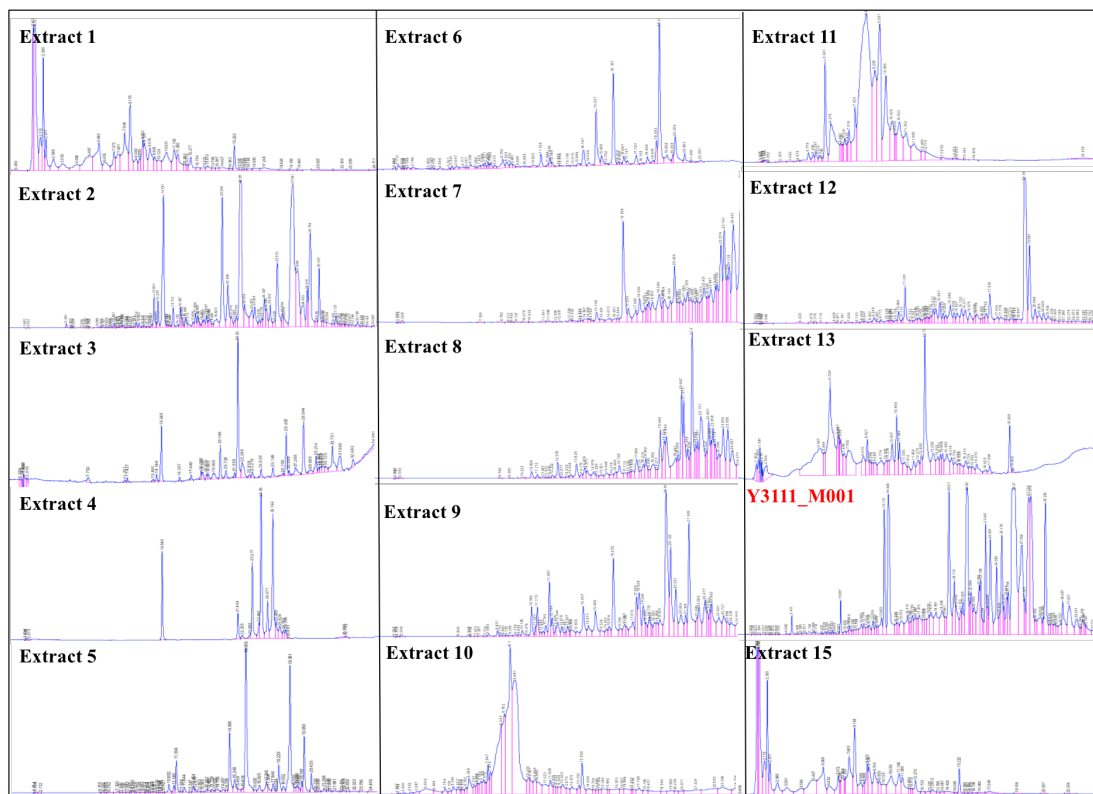
Figure 4.3 HPLC spectra of 15 endophyte crude extracts with the strongest anti-BCG activity.

4.2 Chemical investigation of the secondary metabolites from endophyte Y3111

The strain Y3111, which was isolated from the stems of TCM *Heracleum hemsleyanum*, was identified as a member of *Streptomyces* according to its 16s rRNA gene sequence (Figure 4.4).
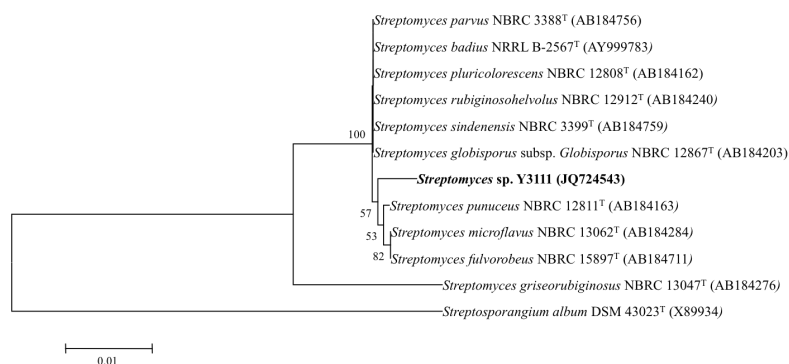
Figure 4.4 16s rRNA gene sequence and neighbor-joining phylogenetic tree of strain Y3111 made by MEGA.

Preliminary analysis of the combined crude extract of Y3111 by HPLC exhibited the presence of several major peaks having a similar UV absorbance profile at 206, 239, 263 and 417 nm (Figure 4.5). The UV absorption maxima at 417 nm suggested the presence of a highly conjugated system. Bioactivity-guided fractionation of the extract led to the isolation of the new compounds (**55–58**) along with 2 known compounds (**59–60**).
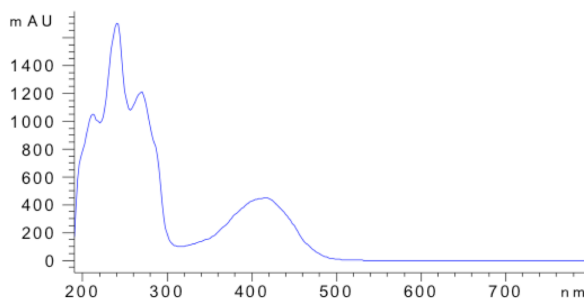
Figure 4.5 UV absorption spectrum of compounds produced by Y3111.

Heraclemycin A (**55**) was isolated as a yellow powder. The molecular formula was deduced as $C_{24}H_{20}O_5$ from its HRESIMS (m/z 389.1421 [M + H]$^+$, calcd. for 389.1389), which possessed 15 degrees of unsaturation. The UV spectrum of heraclemycin A (**55**) showed maximum absorbance at λ_{max} (log ε) 198.0 (5.6), 201.0 (5.6), 206.0 (5.6), 239.0 (5.4), 263.0 (5.1), and 417.0 (4.6) nm in MeOH. The ^{1}H and ^{13}C NMR spectra in
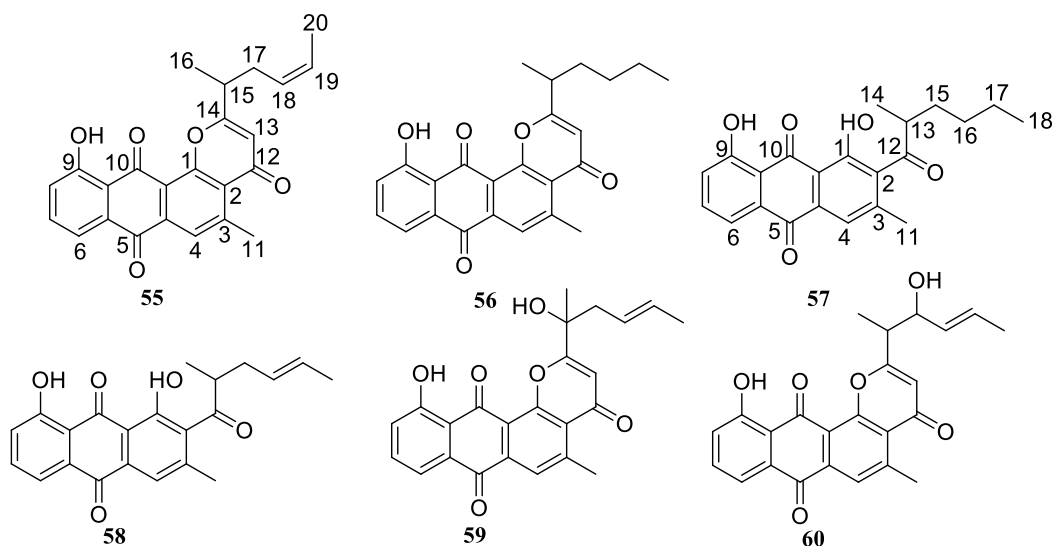
Figure 4.6 Structures of isolated pluramycins (**55**- **60**).

combination with 2D NMR data (COSY, HSQC, and HMBC) (Table 4.1) in CDCl$_3$ revealed the presence of three carbonyl groups (δ_C 187.4, 182.1, and 179.4) conjugated with double bonds, three spin–spin coupling aromatic protons in an ABX system [δ_H/δ_C 7.80 (d, J = 7 .8 Hz)/11 9.4, 7. 66 (t, J = 7.8 Hz)/136.4, and 7.34 (d, J = 7.8 Hz)/125.4], three methyl groups [δ_H/δ_C 1.60 (d, J = 6.7 Hz)/13.0, 1.44 (d, J = 6 .9 Hz)/17 .8, and 3.00 (s)/24 .3], four olefinic carbons, including two isolated olefins [δ_H/δ_C 5.41 (dt, J = 10.8, 8.8Hz)/126.8, 5.54 (dq, J = 10.8, 6.7 Hz)/126.7], and an α,β-unsaturated ketone carbon [δ_H/δ_C 6.23 (s)/111.4 and δ_C 172.6]. One sharp hydrogen resonance at δ_H 12.92 lacked correlation in the HSQC spectrum of heraclemycin A (**55**) (Table 4.1) and was therefore determined to be a hydroxyl group substituted at the aromatic carbon (δ_C 162.7) based on the HMBC correlations of the OH group (δ_H 12.92) to carbons at δ_C 162.7, 125.4 and 117.0. A Dictionary of Natural Products database search using information from combined UV and 1D NMR data indicated the likelihood of a pluramycin skeleton. Heraclemycin A (**55**) exhibited great similarity with the β-indomycinone (**59**), a pluramycin in class of antibiotics sourced from a deep-sea actinomycetete *Streptomyces* sp.,[137] except that the oxygenated carbon in β-indomycinone was replaced by a methine carbon C-15 [δ_H/δ_C 2.85 (m)/39.1]. This indicated that heraclemycin A (**55**) was a new member of the pluramycins and its structure was established as shown in Figure 4.6. All of the signals were assigned unambiguously based on 2D NMR experiments (Table 4.1). The *cis* C-18, C-19 double bond could be determined by the coupling constant between

H-18 and H-19 (J=10.8 Hz).

Table 4.1 ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR data of **55-58** in CDCl$_3$

Position	55		56		57		58	
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)
1	156.8, C		156.9, C		159.8, C		159.7, C	
2	126.6, C		126.7, C		136.4, C		126.4, C	
3	149.8, C		149.8, C		146.4, C		146.5, C	
4	125.6, CH	8.02 s	125.6, CH	8.05 s	122.4, CH	7.70 s	122.4, CH	7.70 s
4a	136.1, C		136.1, C		133.2, C		133.5, C	
5	182.1, C		182.2, C		181.7, C		181.6, C	
5a	132.4, C		132.4, C		133.6, C		133.6, C	
6	119.4, CH	7.80 d (7.8)	119.4, CH	7.81 d (7.8)	120.4, CH	7.85 d (7.8)	120.4, CH	7.84 d (7.2)
7	136.4, CH	7.66 t (7.8)	136.4, CH	7.67 t (7.8)	137.5, CH	7.70 t (7.8)	137.5, CH	7.70 t (7.8)
8	125.4, CH	7.34 d (7.8)	125.4, CH	7.36 d (7.8)	125.0, CH	7.32 d (8.4)	125.0, CH	7.32 d (7.8)
9	162.7, C		162.7, C		162.7, C		162.7, C	
9a	117.0, C		117.0, C		116.0, C		115.5, C	
10	187.4, C		187.4, C		186.3, C		186.8, C	
10a	119.9, C		120.0, C		114.2, C		114.2, C	
11	24.3, CH$_3$	3.00 s	24.4, CH$_3$	3.02 s	20.7, CH$_3$	2.38 s	20.7, CH$_3$	2.38 s
12	179.4, C		179.5, C		209.1, C		206.4, C	
13	111.4, CH	6.23 s	111.3, CH	6.25 s	47.0, CH	3.17 m	47.2, CH	3.23 m
14	172.6, C		173.3, C		15.4, CH$_3$	1.19 d (7.2)	15.3, CH$_3$	1.20 d (7.2)
15a	39.1, CH	2.85 m	39.1, CH	2.79 m	31.8, CH$_2$	1.81 dt (9.7,7.4)	29.2, CH$_2$	2.52 m
15b								2.34 m
16	17.8, CH$_3$	1.44 d (6.9)	18.5, CH$_3$	1.43 d (6.6)	22.9, CH$_2$	1.30-1.36 m	127.5, CH	5.40 dt (17.6,6.6)
17a	31.7, CH$_2$	2.72 dq (14.5,7.3)	34.2, CH$_2$	1.96 m	29.5, CH$_2$	1.30-1.36 m	126.4, CH	5.55 dq (17.6, 6.6)
17b		2.50 dq (14.5,7.3)				1.70 m		
18	126.8, CH	5.41 dt (10.8, 8.8)	29.6, CH$_2$	1.31-1.37 m	14.1, CH$_3$	0.90 t (6.6)	24.9, CH$_3$	1.62 d (6.6)
19	126.7, CH	5.54 dq (10.8, 6.7)	22.7, CH$_2$	1.31-1.37 m	--		--	
20	13.0, CH$_3$	1.60 d (6.7)	14.1, CH$_3$	0.90 t (7.2)	--		--	
1-OH	--		--			12.34 s		12.37 s
9-OH		12.92 s		12.93 s		12.00 s		11.99 s

Heraclemycin B (**56**) was isolated as a yellow powder, and its molecular formula was found to be $C_{24}H_{22}O_5$, which possessed 14 degrees of unsaturation, on the basis of its HRESIMS (m/z 391.1529 $[M + H]^+$, calcd. for 391.1545). The UV spectrum of heraclemycin B (**56**) showed maximal absorbance at λ_{max} (log ε) 199.0 (5.3), 206.0 (5.3), 239.0 (5.1), 267.0 (4.8), and 416.0 (4.3) nm in MeOH. Comparison of ^{1}H NMR and ^{13}C NMR of heraclemycin B (**56**) revealed a high similarity to those of heraclemycin A (**55**), except the absence of the double bond at C-18 and C-19 in heraclemycin B (**56**). As that of heraclemycin A (**55**), the ^{1}H NMR and ^{13}C NMR spectra of heraclemycin B (**56**) showed a characteristic set of signals of the pluramycin class compounds, including: three protons of 1,2,3-trisubstituted benzene ring H-6, H-7, and H-8 $[\delta_H/\delta_C$ 7.81 (d, $J =$ 7.8 Hz)/119.4, 7.67 (t, $J = 7.8$ Hz)/136.4, 7.36 (d, $J = 7.8$ Hz)/125.4], one aromatic proton H-4 $[\delta_H/\delta_C$ 8.05 (s)/125.6], one olefinic H-13 $[\delta_H/\delta_C$ 6.25 (s)/111.32], a hydrogen-bonded phenol proton (δ_H 12.93), and three methyl groups CH$_3$-11, CH$_3$-16, and CH$_3$-20 $[\delta_H/\delta_C$ 3.02 (s)/24.4, 1.43 (d, $J = 6.6$ Hz)/18.5, and 0.90 (t, $J = 7.2$ Hz)/14.1]. Based on the above evidence together with the 2D data of heraclemycin B (**56**) (Table 4.1), the structure of heraclemycin B (**56**) could be determined (Figure 4.6).

Heraclemycin C (**57**) was obtained as a yellow powder. The molecular formula was determined by HRESIMS as $C_{22}H_{22}O_5$ (m/z 367.1568 $[M + H]^+$, calcd. for 367.1501), requiring 12 degrees of unsaturation. The ^{1}H NMR spectrum of heraclemycin C (**57**) contained resonances for two phenolic hydroxyl hydrogen $[\delta_H$ 12.00 and 12.34]. Comparison of the ^{1}H and ^{13}C NMR spectra of heraclemycin C (**57**) with those of heraclemycin B (**56**) revealed that both shared a similar anthraquinone skeleton, and the significant differences indicated the reduction of the γ-pyrone moiety, culminating in the absence of the ethenyl group in the ^{1}H and ^{13}C NMR spectrum and the presence of an unconjugated carbonyl functionality $[\delta_C$ 209.1]. The location of the hydroxyl group $[\delta_H$ 12.34] at C-1 was corroborated by the HMBC correlation form 1-OH to C-2 $[\delta_C$ 136.4], C-10a $[\delta_C$ 114.2] and C-1 $[\delta_C$ 159.8]. All the signals of heraclemycin C (**57**) were assigned unambiguously based on 2D NMR experiments (Table 4.1) and its structure was determined as shown in Figure 4.6.

Heraclemycin D (**58**) was obtained as a yellow powder and the HRESIMS data suggested the molecular formula $C_{22}H_{20}O_5$ (m/z 365.1440 $[M + H]^+$, calcd. for 365.1344). In the ^{1}H

NMR spectrum signals of heraclemycin D (**58**) were very similar to those of heraclemycin C (**57**) except for the presence of double bond at C-16 and C-17 in 4 [δ_H/δ_C 5.40 (dt, J = 17.6, 6.6 Hz)/127.5, 5.55 (dq, J = 17.6, 6.6 Hz)/126.4]. The trans C-18, C-19 double bond could be determined by the J value (17.6 Hz) between H-16 and H-17. Thus, the structure of heraclemycin D (**58**) could be determined and all the NMR signals were assigned unambiguously based on 2D NMR experiments (Table 4.1).

All of the isolated compounds were evaluated for their biological activities against *M. bovis* BCG, *S. aureus*, MRSA, *Bacillus subtilis*, and *Pseudomonas aeruginosa* from 100 μg/mL and MICs were calculated to μM (Table 4.2). Heraclemycin C (**57**) and heraclemycin D (**58**) showed significant activity against BCG with the MIC values of 17 and 62 μM, respectively. Heraclemycin A (**55**) showed weak activity with an MIC value of 258 μM, whereas heraclemycin B (**56**), β-indomycinone (**59**) and saptomycin A (**60**) were inactive against BCG. These results indicated that the ring-closed pyrone resulted in loss of activity, and the ring-opened form was moderately active. On the other hand, compounds heraclemycin A (**55**) and D (**58**) exhibited anti-Gram positive bacteria activity including anti-BS, anti-SA, and anti-MRSA. However, compounds heraclemycin B (**56**), C (**57**), β-indomycinone (**59**) and saptomycin A (**60**) were unable to inhibit these bacteria up to 100 μg/mL, suggesting that hexenyl side chain at C-13 or C-15 is important for these microbe inhibitory activities.

Table 4.2 Antimicrobial activities of compounds

Organism (Strain)	Minimum Inhibitory Concentration (μM)						
	55	56	57	58	59	60	Control
Bacillus Calmette-Guérin (Pasteur 1173P2, BCG)	258	NA	17	62	NA	NA	0.37[a]
Staphylococcus aureus (ATCC 6538)	129	128	NA	62	NA	NA	0.7[b]
Methicillin-resistant *S. aureus* (Clinical strain of Chaoyang hospital)	129	128	NA	62	NA	NA	0.7[b]
Bacillus subtilis (ATCC 6633)	258	NA	NA	124	NA	NA	0.35[b]
Pseudomonas aeruginosa (PAO1)	NA*	NA	NA	NA	NA	NA	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

Time-kill experiments of the new pluramycin-type compounds against BCG were performed. Time-kill analyses were performed according to CLSI method M26-A.20

(1999). Antimicrobials were considered bactericidal at the lowest concentration that reduced the original inoculum by ≥3 $\log_{10}$ RFU/mL (99.9 %) at each of the time periods and bacteriostatic if the inoculum was reduced by 0 to 3 $\log_{10}$ RFU/mL.[138] As shown in Figure 4.7, maximum killing for heraclemycin C was observed at concentrations of 8× the MIC, with a 1.2-log drop in the numbers of RFU/mL occurring by 108 h after the addition of the compound, consistent with a bacteriostatic effect (Figure 4.7C). Just slightly less than 1.2-log drop in RFU/mL was also observed for heraclemycin D at 2× and 4× the MIC (Figure 4.7D). However, even in the presence of 100 μg/mL heraclemycin A and B, there were no obvious differences between compounds treatment and control, suggesting an inactive effect on mycobacteria (Figure 4.7A, B).
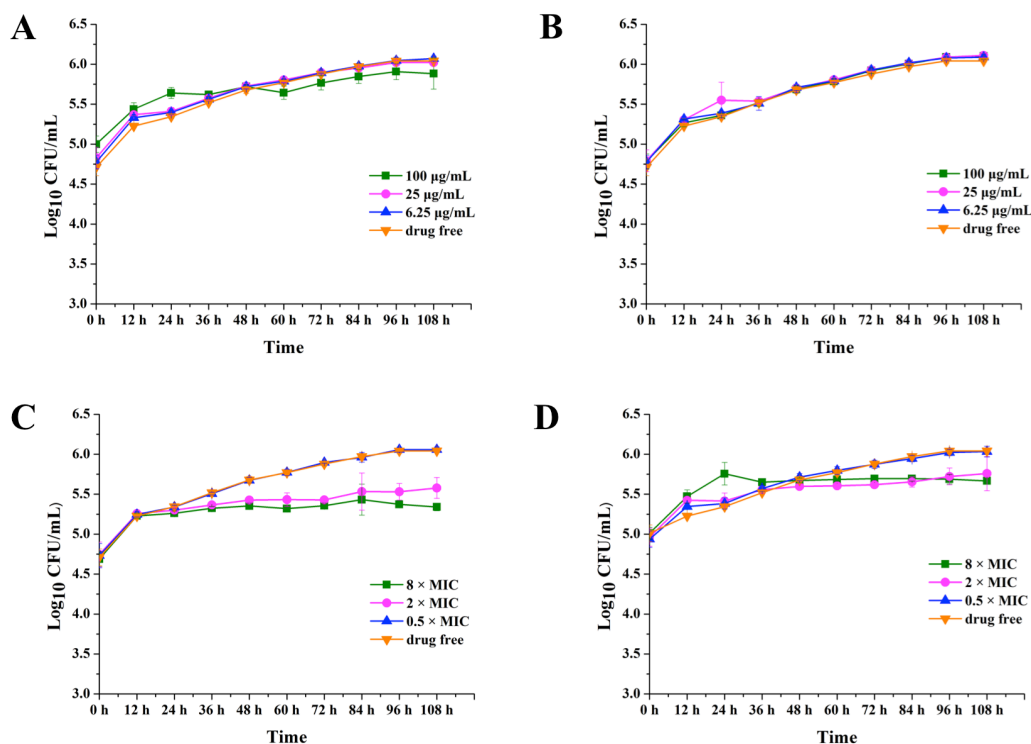
Figure 4.7 Killing activities of compounds 55– 58 against M. bovis BCG. Compounds were added to cultures at time zero. A. heraclemycin A (**55**), B. heraclemycin B (**56**), C. heraclemycin C (**57**), D. heraclemycin D (**58**).

4.3 Conclusion

In our present investigation, we identified the first example of pluramycin analogues produced by TCM endophytic microbes and the first BCG-selective pluramycin which

indicates a potentially unique mode of action. Previously reported pluramycin antibiotics containing the 4H-anthra [1,2-*b*] pyran-4,7,12-trione nucleus to which amino sugars such as angolosamine and vancosamine are typically attached by C -glycosidic linkage at C-6 and C-8 were found to have versatile and strong antimicrobial and anticancer activities.[139] Compared to other natural sources, endophytes are a poorly investigated group of microorganisms that represent an abundant and dependable source of bioactive and chemically novel compounds with potential for exploitation in a wide variety of medical, agricultural, and industrial arenas.[140] Our investigations confirm that screening of endophytes offers a valuable approach towards the discovery of new anti-TB natural products. In addition, our study extends knowledge of the secondary metabolism of an endophyte associated with TCM *H. souliei* (Y3111). Bioassay-guided fractionation yieded the structurally diverse heraclemycins A–D, and in particular heraclemycin C, which showed strong anti-BCG activity with a MIC value of 6.25 μg/mL.

In numerous cases, the bioactivity-guided purification of individual constituents results in a considerable work taking several days even several weeks or months is still necessary to isolate and elucidate individual metabolite structures from crude natural extracts. This procedure has been developed to focus only on the fractions or metabolites with a defined biological activity. However, often such approaches are applied to finally rediscover already known compounds or even end with the loss of activity. The potential new and active natural products present at minor or trace levels can easily be overlooked.

In view of these observations, a more specific, rapid and sensitive approach to chemical characterization is needed.

Chapter V NMR fingerprinting-guided investigation of microbial crude extracts from unique environments

Abstract: This chapter presents an investigation for using a strategy of NMR fingerprinting instead of bioactivity to guide the identification of natural products from small-scaled microbial extracts. Lead-like enhanced fractions were generated from 50 microbial crude extracts. Then 550 fractions were screened against BCG and 142 of them were active. NMR-guided isolation and purification gave 25 natural products, including 11 potentially active compounds.

5.1 Introduction

Microorganisms collected from marine, plants and desert appear to have adapted to their extreme environments, thus gaining their diverse metabolic and genetic capabilities. With more and more chemically interesting and biologically significant metabolites isolated from these microorganisms, it has been shown as a rich source of novel compounds with various biological activities. Among the microorganisms, the actinomycetetes group continues to be a prolific source of bioactive compounds. In particular, although the exploitation of marine actinomycetetes as a source for the discovery of novel secondary metabolites is still at an early stage, the discovery rate of novel secondary metabolites from marine actinomycetetes has recently surpassed that of their terrestrial counterparts.[141-143] As previously mentioned, a cell-based screening assay was used to identify inhibitors in an aerobic, logarithmic growth screen of *Mycobacterium bovis* BCG from a natural product library. In the course of the screening program for anti-BCG agents from the NPL, small-scale extracts of a batch of actinomycetetes strains isolated from the South China Sea, Traditional Chinese Medicines and the Taklimakam desert were tested. A subset of the screening hits, namely 45 strains resulting in 64 crude extracts were selected for further investigation.
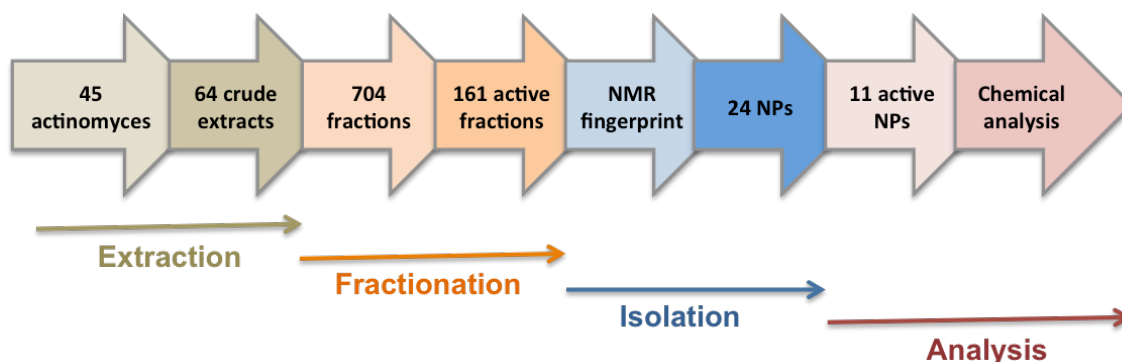
Figure 5.1 An overview of the natural product discovery program.

The traditional process of discovering new bioactive natural products via bioassay-guided isolation is generally long and laborious, and known natural products are frequently rediscovered. Therefore, the main purpose of this work is to establish a strategy for rapid identification of natural products by a combination of HPLC fractionation and ^{1}H NMR fingerprint technique. In this chapter, we describe the method development and by applying the method, chemical and biological properties of the selected crude extracts

were investigated. The overall outline of the natural product discovery program is shown in Figure 5.1.

5.2 Lead-like enhanced (LLE) gradient optimization

The lead-like enhanced (LLE) gradient was developed to collect the fractions that are most likely to contain compounds with drug-like properties.[144] However, according to our research, active compounds derived from marine microbes or endophytes can be rather non-polar, which are not collected by the traditional LLE gradient that collects fractions from 2 to 7 minutes. Shift the start of fraction collection may result in missing out the compounds with mid polarity. Therefore, to increase the probability of quality hits of marine microbes and endophytes, we did a LLE gradient optimization experiment.

As shown in Figure 5.2, two new gradients (gradient 2 and 3) with longer methanol-eluting times than the original gradient were developed. In order to find the best HPLC conditions for finding TB drug-like molecules, four crude extracts, which produced non-polar compounds, were selected to validate the 3 gradients.
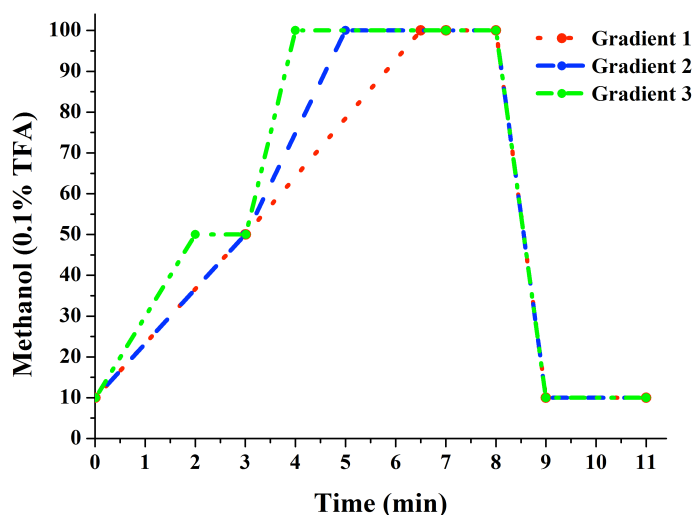
Figure 5.2 HPLC conditions of gradient 1-3. Solvent A: 0.1% TFA in methanol, solvent B: 0.1% TFA in water; flow rate: 4 mL/min from 0-3 min and 7-11 min, 3 mL/min from 3-7 min.

Four test samples (5 mg) were dissolved in 500 μL DMSO. Then the DMSO solution (100 μL) of each extract was fractionated by C_{18} analytical HPLC with gradients 1-3. The HPLC traces are shown in Figure 5.3. Despite the fact that some of the non-polar constituents of the test samples were included in the collected fractions using gradient 2, gradient 3 showed more impressive improvement in the inclusion of non-polar molecules. According to the result of this LLE gradient optimization, gradient 3 was selected as the

new gradient for generating LLE fractions.
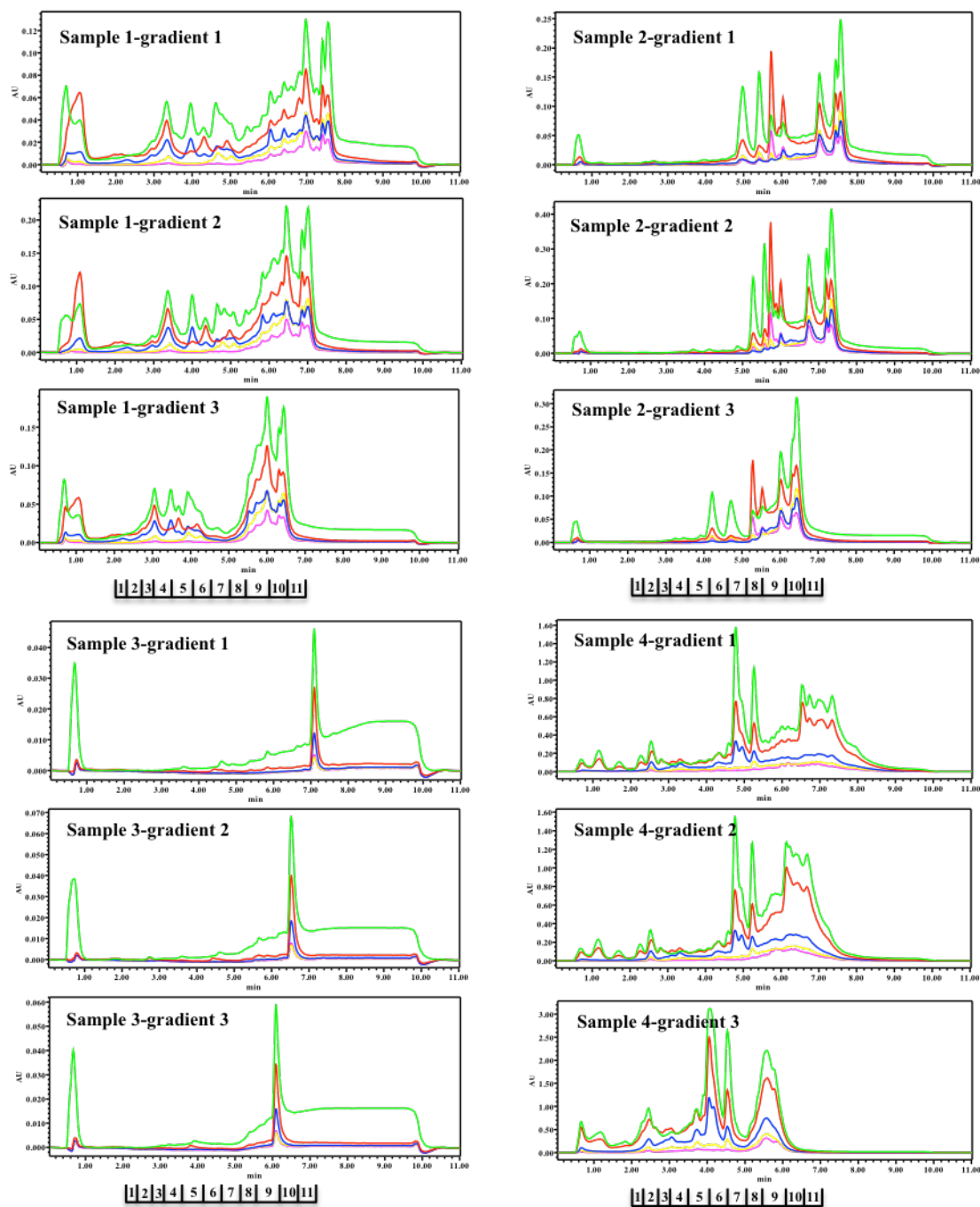
Figure 5.3 The HPLC traces of the four test samples fractionated by gradients 1-3.

5.3 Fraction library screening and ^{1}H NMR fingerprints of selected extracts and fractions

The 64 crude EtOAc extracts from 45 TB-active actinomycetes, including 4 endophytes, 16 desert strains and 30 marine strains, were dissolved in DMSO at the concentration of

10 $\mu g/\mu L$. The fraction library was constructed using reverse-phase solvent conditions (MeOH/H_2O/0.1% TFA) on a C_{18} Phenomenex Onyx monolithic HPLC column. Eleven fractions were collected per extract between 2 and 7 min of the run. The fractionation process also separated the complex crude extracts into fractions containing a small number of compounds to facilitate the rapid identification of active molecules. In total, the fraction library was composed of 704 fractions from 64 extracts. The 704 lead-like enhanced fractions were subsequently screened and 161 hits (22.87%, single dose) showed the activity (Figure 5.4).
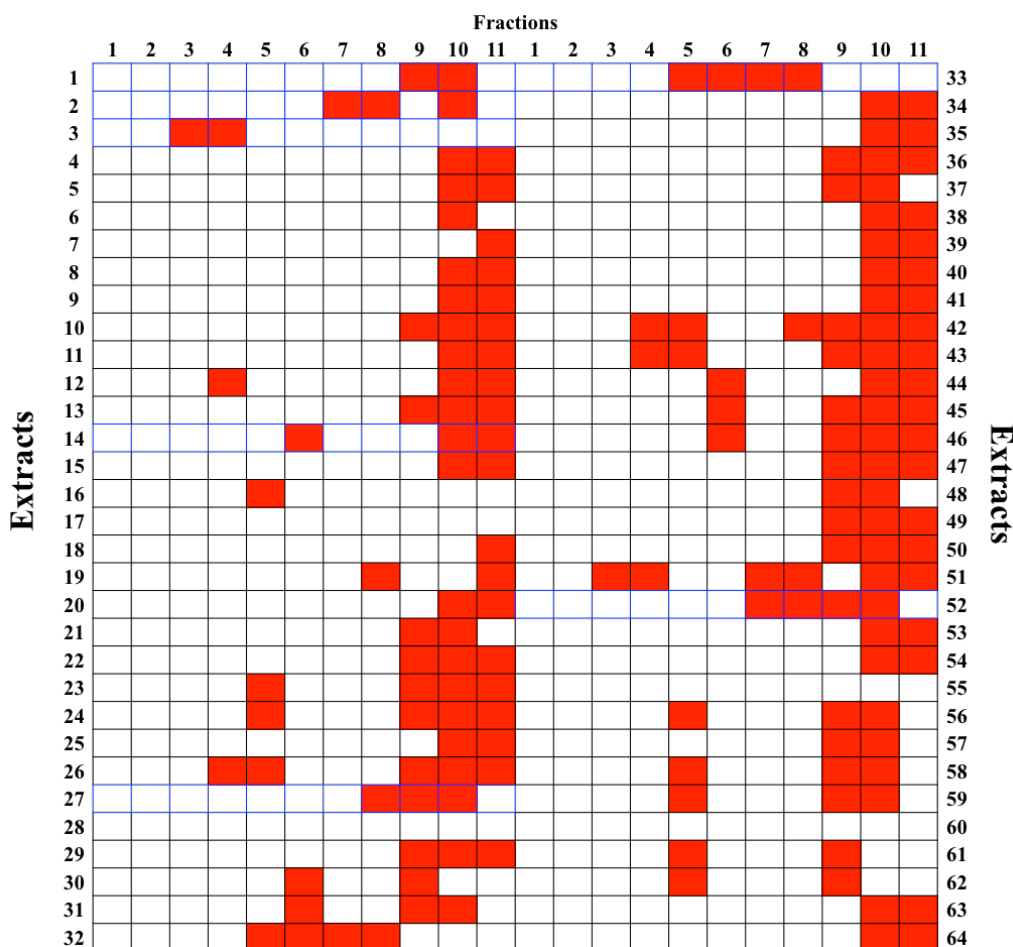
Figure 5.4 Overview of the anti-BCG screening on LLE fractions. There are 704 fractions prepared from 64 crude extracts, blanks filled with red color represents the BCG activity. 8 selected crude extracts are in blue blanks.

The 161 active fractions belonged to 42 different active samples while 3 of the active extracts lost their activities after fractionation. Further analysis showed that the fractions were distributed across fraction 3 to fraction 11, with 72.7% originating from the

relatively non-polar fractions 9-11, confirming the importance of gradient optimization.

5.4 Isolation and identification of natural products from selected extracts

Based on confirmed activity by screening assay testing on BCG and amount of material available, 21 out of 161 prioritized fractions, representing 7 strains (fractions 9 and 10 of ES120055_AM2, fractions 7, 8 and 10 of ES120127_AM2, fractions 3 and 4 of ES130159_M001, fractions 6, 10 and 11 of LS120167_AM2, fractions 8, 9 and 10 of MS110105_MPG, fractions 5 to 8 of MA110109_MPG and fractions 7 to 10 from MS110154_MPG), were initially chosen for NMR fingerprint evaluation and small-scaled isolation work (Figure 5.4). The standard VnmrJ 3.2 Proton pulse sequence was run with the following parameters: pw = 45°, p1 = 0 μs, d2 = 0 s, d1 = 1s, at = 1.7 s, sw = 9615Hz, nt = 128 scans.

5.4.1 Small-scale isolation of MS110154_MPG

Marine actinomycete MS110154 was isolated from a sediment sample collected from the South China Sea. It was identified as a *Streptomyces* sp. based on cultural, physiological, morphological characteristics and 16s rRNA gene analysis (Figure 5.5).

Figure 5.5 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110154 and related species constructed by MEGA.

Despite the fact that fraction 7 showed weak signals in NMR fingerprint, fractions 8, 9 and 10 of MS110154_MPG presented interesting proton signals in aromatic region and sugar region as well, especially fraction 9 (Figure 5.6).

NMR-guided small-scale isolation led to the identification of one new compound (**61**) as well as two known compounds chrysomycins A (**62**) and B (**63**) (Figure 5.7).
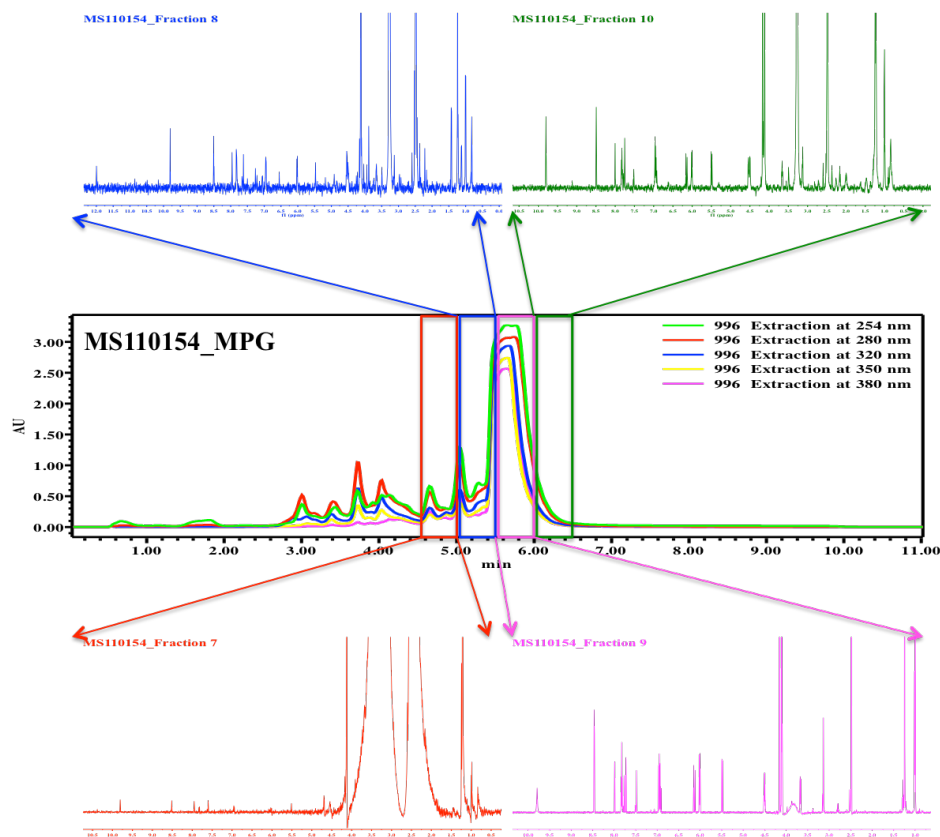
Figure 5.6 HPLC trace and NMR fingerprint of MS110154_MPG.
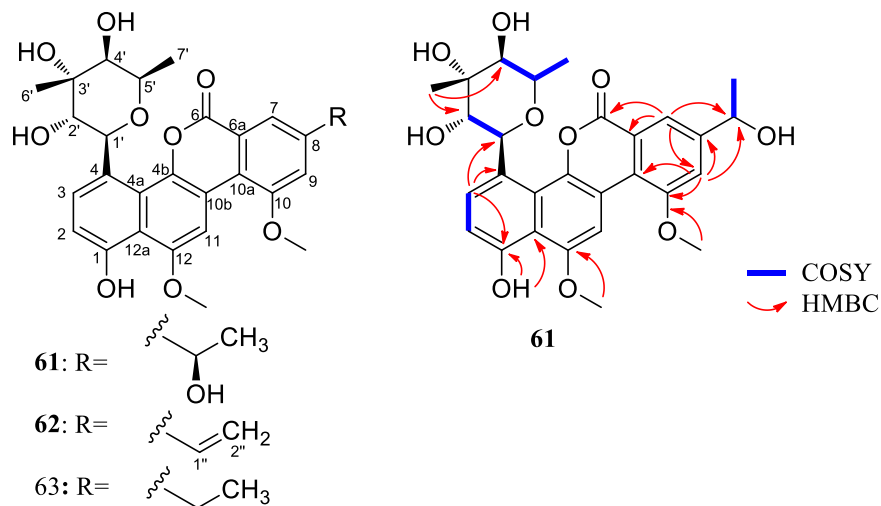
Figure 5.7 Structures and observed COSY and key HMBC correlations of **61 – 63**.

The HPLC chromatogram (Figure 5.8) displayed two major peaks, all with similar UV spectra (λ_{max} 245, 270, 308 and 385nm), indicating that both compounds belong to the same chemical class. The LC–MS chromatograms displayed masses m/z 509 and 497 [M + H]$^+$ for peaks appearing at t_R 8.89 and 8.61 min, respectively. The ^{1}H NMR

experiments for two compounds in DMSO-d_6 were then conducted and a benzo-naphtha-pyran skeleton was found in both compounds. The information related to the proton NMR and UV profiles of the two compounds was then used for the search in the DNP for rapid dereplication. This search returned two classes of antibiotics gilvocarcins and chrysomycins. Filtered by applying the mass range, compound **62** was identified as chrysomycin A and **63** was identified as chrysomycin B. The structures were also confirmed by comparison of their NMR with literature values and the correlations in 2D NMR (Table 5.1).
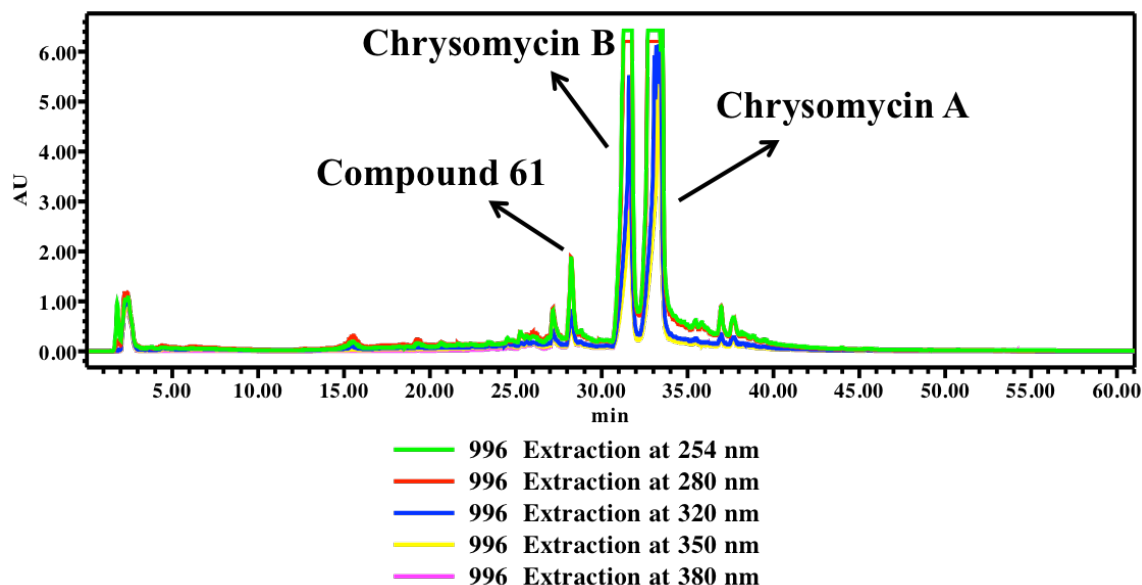
Figure 5.8 HPLC chromatogram of MS110154_MPG and isolated compounds.

Compound **61** was obtained as yellow powder with the molecular formula of $C_{28}H_{30}O_{10}$ from the LC-MS peak at m/z 527 [M+H]$^+$. Its UV spectrum showed characteristic peaks of chrysomycin-type polyketide chromophore at λ_{max} 245.0, 269.9, 307.7 and 384.9 nm. The ^{1}H NMR spectrum showed 30 proton signals including five aromatic protons at δ_H 6.95 ~ 8.50, two methoxy groups, three C-methyl groups and five oxygenated methine protons at δ_H 3.13 ~ 6.02 ppm (Table 5.2). The ^{13}C NMR spectrum displayed 28 carbon signals that were classified by HSQC and HMBC spectra as two methoxy carbons, three methyl carbons, five olefinic methine carbons, five oxygenated methane carbons, 11 fully substituted olefinic carbons, 4 of which are oxygenated, and one conjugated ester carbonyl carbon.

Table 5.1 ^{1}H and ^{13}C (600 MHz) NMR data of **62** and **63** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC	δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC
	62				**63**			
1	153.3 C				153.4 C			
2	112.5 CH	6.96 d (8.4)	3	4, 12a	112.2 CH	6.94 d (8.4)	3	4, 12a
3	129.5 CH	7.83 d (8.4)	2	1, 4a 1'	129.6 CH	7.82 d (8.7)	2	1, 4a 1'
4	128.7 C				128.3 C			
4a	125.7 C				125.4 C			
4b	142.8 C				142.3 C			
6	160.2 C				159.9 C			
6a	117.6 C				122.2 C			
7	119.4 CH	7.99 d (1.3)		6, 6a, 9, 10a	121.6 CH	7.78 d (1.4)		6, 6a, 9, 10a, 1"
8	139.1 C				140.8 C			
9	115.2 CH	7.74 d (1.3)		7, 10, 10a, 1"	119.5 CH	7.50 d (1.6)		7, 8, 10, 10a, 1"
10	157.9 C				157.3 C			
10a	123.4 C				122.3 C			
10b	113.4 C				113.6 C			
11	101.8 CH	8.48 s		4b, 10b, 12, 12a	102.2 CH	8.48 s		4b, 10b, 12, 12a
12	152.1 C				152.1 C			
12a	114.9 C				115.1 C			
1'	75.1 CH	6.01 d (9.3)	2'	3, 2'	74.8 CH	6.01 d (9.6)	2'	3, 2'
2'	72.9 CH	3.66 t (8.4)	1', 2'-OH	1'	73.0 CH	3.67 t (8.9)	1', 2'-OH	1'
3'	74.2 C				74.3 C			
4'	76.3 CH	3.13 d (7.5)	4'-OH	2', 6'	76.4 CH	3.13 d (7.6)	4'-OH	2', 6'
5'	70.8 CH	4.50 dd (6.0, 7.3)	6'		71.0 CH	4.50 dd (6.3, 7.1)	6'	
6'	24.1 CH$_3$	1.24 s		2', 4'	24.2 CH$_3$	1.24 s		2', 4'
7'	17.3 CH$_3$	1.00 d (6.3)	5'	4', 5'	17.2 CH$_3$	1.00 d (6.5)	5'	4', 5'
1"	135.4 CH	6.92 dd (11.5, 17.6)	2"a, 2"b	7, 8, 9, 2"	21.5 CH$_3$	2.50 s		7, 8, 9
2"	21.5 CH$_2$	6.13 d (17.3)	1", 2"b	8, 1"				
		5.49 d (11.6)	1", 2"a	8, 1"				
1-OH		9.80 s		1, 2, 12a		9.80 s		1, 2, 12a
2'-OH		4.11 m	2'			4.13 d (8.3)	2'	
3'-OH		4.16 s		2', 4'		4.16 s		2', 4'
4'-OH		4.54 d (8.2)	4'			4.53 d (7.7)	4'	
10-OCH$_3$	57.1 CH$_3$	4.16 s		10	56.8 CH$_3$	4.13 s		10
12-OCH$_3$	56.6 CH$_3$	4.11 s		12	56.8 CH$_3$	4.12 s		12

Table 5.2 ^{1}H and ^{13}C (600 MHz) NMR data of **61** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC	ROESY
1	153.7 C				
2	112.4 CH	6.95 d (8.3)	3	4, 12a	3
3	129.7 CH	7.82 d (7.9)	2	1, 4a 1'	2, 2'
4	128.4 C				
4a	125.4 C				
4b	142.7 C				
6	160.4 C				
6a	122.4 C				
7	115.6 CH	7.61 d (1.7)		6, 6a, 9, 10a, 1"	1", 2", 10-OCH$_3$
8	150.2 C				
9	118.1 CH	7.96 d (1.3)		7, 8, 10, 10a, 1"	1", 2"
10	157.4 C				
10a	122.2 C				
10b	113.3 C				
11	102.0 CH	8.50 s		4b, 10a, 10b, 12, 12a	12-OCH$_3$
12	152.3 C				
12a	115.7 C				
1'	74.9 CH	6.02 d (9.3)	2'	3	2', 5', 3'-OH
2'	73.1 CH	3.66 t (8.1)	1', 2'-OH	1'	3, 1', 7', 3'-OH, 4'-OH
3'	73.4 C				
4'	76.3 CH	3.13 d (7.8)	4'-OH	2', 4'	5', 6', 7', 3'-OH, 4'-OH
5'	71.4 CH	4.50 dd (6.8, 4.3)	6'		1', 4', 6'
6'	24.3 CH$_3$	1.24 s		2'	2', 4', 3'-OH, 4'-OH
7'	17.5 CH$_3$	1.00 d (6.3)	5'	4', 5'	4', 5', 4'-OH
1"	68.0 CH	4.92 t (6.0)	1"-OH, 2"		7, 9, 2"
2"	26.0 CH$_3$	1.42 d (6.4)	1"	8', 1"	7, 9, 1"
1-OH		9.80 s		1, 2, 12a	12-OCH$_3$
2'-OH		4.13 m	2'		
3'-OH		4.16 s		2', 4'	1', 2', 4', 7'
4'-OH		4.54 d (9.1)	4'		2', 6', 7', 2"
1"-OH		5.48 d (4.9)	1"		
10-OCH$_3$	57.0 CH$_3$	4.13 s		10	7
12-OCH$_3$	57.0 CH$_3$	4.12 s		12	11, 1-OH

These NMR data were very similar to those of chrysomycin A further indicating the nature of chrysomycin-type polyketide glycoside of **61**. Compared with those of chrysomycin A, the NMR differences of **61** was captured in the replacement of $-CH=CH_2$ signals by $-CH(OH)CH_3$ signals that was further supported by $^1H-^1H$ COSY correlations between H$_3$-2″ (δ_H 1.42) with H-1″ (δ_H 4.92) and H-1″ with OH-1″ (δ_H 5.48). Furthermore, obvious downfield shifts for C-8, H-7 and H-9 and upfield shifts for C-7 were observed between **61** and chrysomycin A. These observations suggested that compound **61** is the hydrated derivative of chrysomycin A at the C-8 side chain, which was further supported by the key HMBC correlations from H-7 (δ_H 7.61, d, J=1.3 Hz) to C-6 (δ_C 160.4), C-6a (δ_C 122.4), C-9 (δ_C 115.6) and C-1″ (δ_C 68.0), from H-9 (δ_H 7.96, d, J=1.7 Hz) to C-8 (δ_C 150.2), C-10 (δ_C 157.4), C-10a (δ_C 122.2) and C-1″, and from H$_3$-2″ to C-8 (Table 5.2).

To investigate the biological effects of the compounds, we tested the anti-BCG, anti-SA, anti-MRSA, anti-BS and anti-PA activities (Table 5.3). Of the 3 compounds, 2 known compounds displayed potency to different extents against the growth inhibition of Gram-positive bacteria. Chrysomycins A (**62**) was more active with MIC of 6 μM for BCG, SA and MRSA and <1.5 μM for BS. However, compound **61** did not show significant effect on any of the bacteria at the tested concentration, indicating that the size of side chain attached in C-8 is crucial for activity.

Table 5.3 Antimicrobial activities of compounds **61 – 63**

Organism (Strain)	Minimum Inhibitory Concentration (μM)			
	61	**62**	**63**	Control
Bacillus Calmette-Guérin (Pasteur 1173P2, BCG)	NA*	6	25	0.37[a]
Staphylococcus aureus (ATCC 6538)	NA	6	50	0.7[b]
Methicillin-resistant *S. aureus* (Clinical strain of Chaoyang hospital)	NA	6	50	0.7[b]
Bacillus subtilis (ATCC 6633)	NA	1.5	25	0.35[b]
Pseudomonas aeruginosa (PAO1)	NA	NA	NA	3[c]

[a] isonizid [b] vancomycin [c] ciprofloxacin * NA: not active

Almost all the NMR signals in active fractions of MS110154 can be found in the responding compounds, except for a few weak signals in fraction 8 (Figure 5.9).

Figure 5.9 Comparison of NMR fingerprint between active fractions and compounds of
MS110154_MPG

5.4.2 Small-scale isolation of LS120167_AM2

LS120167 was isolated from a desert sample collected from Taklimakan Desert and was
identified as a *Streptomyces* sp. based on 16s rRNA phylogenetic analysis (Figure 5.10).
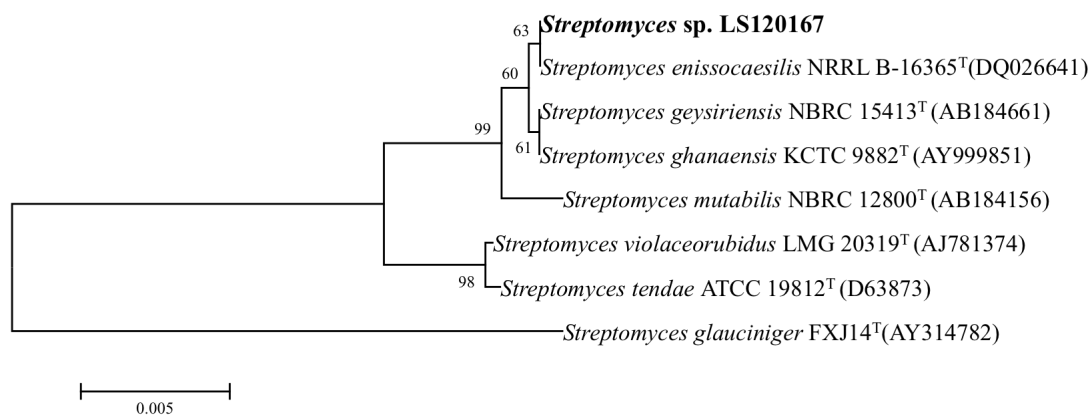
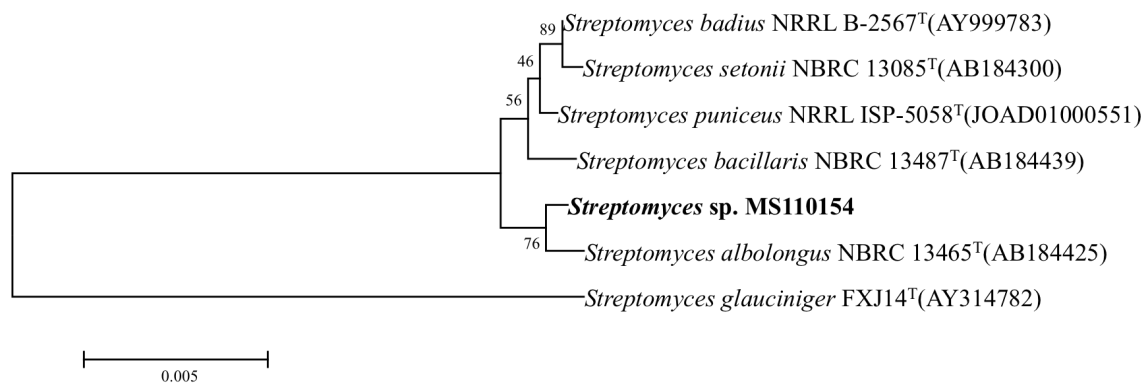Figure 5.10 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of LS120167 and
related species constructed by MEGA.

BCG-based activity test of the 11 fractions of LS120167_AM2 revealed that fractions 6, 10 and 11 showed positive results. Fraction 6 didn't display much information in its proton spectrum except for several methyl groups, however, fingerprint of fraction 10 and 11 showed different structure pattern with fraction 6, such as methyl and unsaturated protons, indicating some interesting molecules were produced by this strain (Figure 5.11).
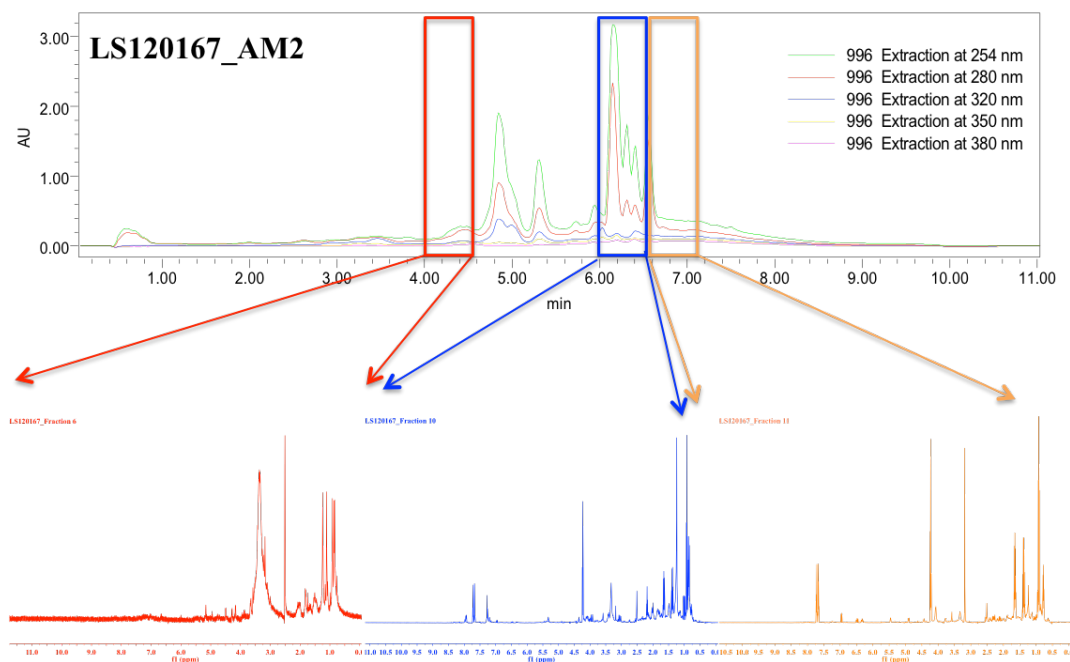
Figure 5.11 HPLC trace and NMR fingerprint of LS120167_AM2.

Two active components were isolated from fraction 10 and 11. Borrelidin (**64**) was isolated as a yellow solid (Figure 5.12). The molecular weight of the compound was ascertained by LC-MS analysis and the molecular ion peak was found at 489.5. ^{13}C NMR confirmed the presence of twenty-eight carbons and ^{1}H NMR about the presence of forty-three protons (Table 5.4). A carboxyl group with chemical shift of 171.9 ppm was also determined from ^{13}C NMR spectrum. Furthermore, four methyl carbons were found in the range δ values of 14-20 ppm in ^{13}C NMR, while their corresponding protons were found in δ values of 0.8-1.1 ppm in ^{1}H NMR. Similarly, the conjugated diene was identified by the presence of signals in δ_C 127-145 ppm in ^{13}C NMR and their corresponding protons in δ values of 6.3-7.0 ppm in ^{1}H NMR. The NMR data was compared with reported data of borrelidin and found to be similar.[145] The absolute configuration of **64** was determined by comparing the $[\alpha]_D$ value in ethanol with literature.[146]

Table 5.4 ^{1}H and ^{13}C (600 MHz) NMR data of **64** in CD$_3$OD

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1	171.9, C			
2	36.7, CH$_2$	2.25 dd (15.0, 10.3)	2', 3	1, 15
		2.40 dd (16.0, 10.3)	2, 3	1, 2
3	71.6, CH	3.94 d (9.7)	2, 2', 3	1, 2, 5, 24
4	35.7, CH	1.83 m	3, 5, 24	2
5	43.2, CH$_2$	1.22 m	4, 6	3, 4, 6, 7, 24, 25
		0.99 t (12.2)	4, 6	3, 4, 5, 6, 7, 24, 25
6	27.2, CH	1.83 m	5, 7, 25	8
7	26.3, CH$_2$	1.59-1.71 m	6, 8	4, 8
8	48.0, CH	0.99 t (12.2)	7, 9, 26	6, 8
9	37.7, CH$_2$	1.22 m	8, 9', 10	7, 8, 10, 11, 26, 27
		0.72 t (12.0)	8, 9, 10	7, 8, 10, 26, 27
10	34.6, CH	1.83 m	9, 11, 27	11, 12
11	71.5, CH	4.20 d (9.5)	10	9, 10, 12, 13, 27, 28
12	116.1, C			
13	144.0, CH	6.92 d (11.2)	14	11, 12, 14, 15, 28
14	127.6, CH	6.61 dd (14.5, 12)	13, 15	12, 13, 16
15	138.7, CH	6.33 ddd (13.2, 9.6, 3.7)	14, 16	13, 16, 17
16	35.3, CH$_2$	2.50-2.63 m	15, 17	14, 15, 18
17	76.0, CH	5.00 d (5.0)	16, 18	1, 15, 19, 22
18	46.1, CH	2.69 p (8.9)	17, 19, 22	16, 19, 22, 23
19	29.1, CH$_2$	2.01 m	18, 19', 20	18, 20, 21, 22
		1.43 m	18, 19, 20	18, 20, 21
20	24.8, CH$_2$	1.83 m	19, 21, 21'	18, 19, 21, 22
21	31.1, CH$_2$	2.01 m	20, 21', 22	18, 19, 20, 22, 23
		1.83 m	20, 21, 22	18, 19, 20, 22, 23
22	48.6, CH	2.46 q (7.4)	18, 21, 21'	18, 20, 21, 23
23	178.8, C			
24	17.4, CH$_3$	0.87 d (6.3)	4	4, 5
25	19.6, CH$_3$	0.87 d (6.3)	6	5, 6
26	17.6, CH$_3$	0.87 d (6.3)	8	9
27	14.1, CH$_3$	1.05 d (5.4)	10	9, 10
28	118.5, C			
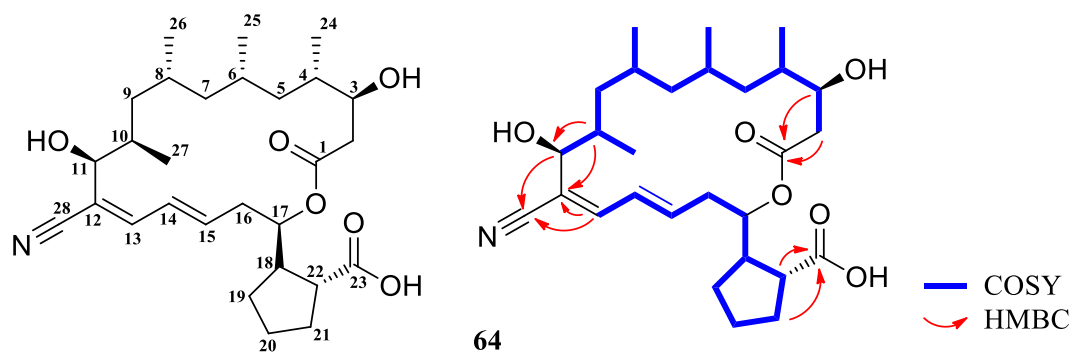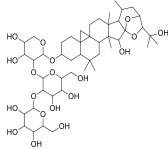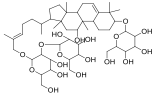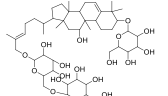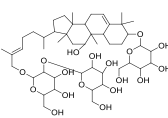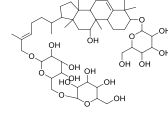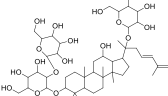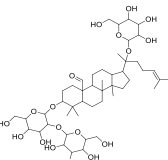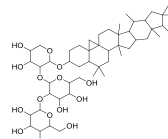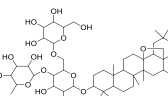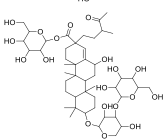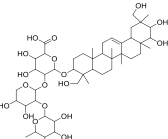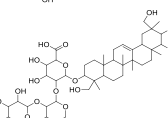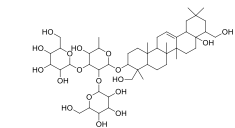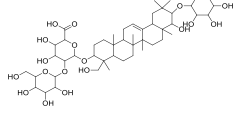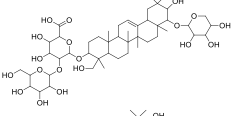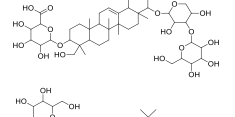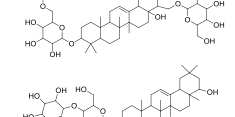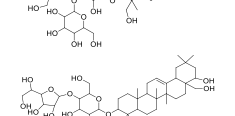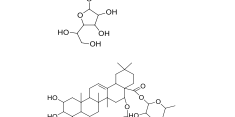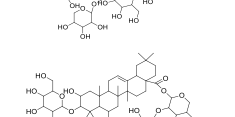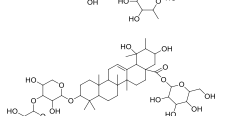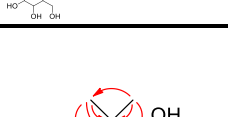
Figure 5.12 Structures and observed COSY and key HMBC correlations of borrelidin (**64**)

Kudzusaponin SA2 (**65**) was isolated as a colorless powder. Dereplication of this compound proved to be challenging and structure elucidation work had to be undertaken. Its molecular weight was determined from the data of HRESIMS (m/z 945.5055, [M+H]$^+$). ^{13}C NMR showed the presence of 47 carbons and ^{1}H NMR about the presence of 76 protons (Table 5.6). They comprised of an olefinic proton at δ_H 5.20 (t, J=3.5 Hz, H-12), seven singlets (each 3H) were characteristic of the ursene skeleton. Moreover, proton chemical shifts between 3 and 5 ppm were found to be typical signals of sugars. Then a dereplication work was done by searching in DNP and 22 compounds were found (Table 5.5), which had the same molecular weight and possessed a triterpene skeleton as well as several sugar substitutes. After comparing the numbers of methyl and methylene, which were determined by HSQC NMR (seven methyl and eleven methylene, Table 5.5) between **65** and each compound in Table 5.14, only three compounds were left, namely lupinoside PA2 (NO. 14), kudzusaponin SA2 (NO. 15) and impatienoside B (NO. 16). HMBC correlations from H-22 (δ 3.33) to C-1‴ (δ 107.4); from H-3 (δ 3.42) to C-1′ (δ 103.5); from H-2′ (δ 3.60) to C-1″ (δ 103.5) established the connection of terterpene skeleton and three sugar units and then the structure of **65** was determined as kudzusaponin SA2 (Figure 5.13). The absolute configuration of **65** was confirmed by compared the [α]$_D$ value in methanol with literature.[147]
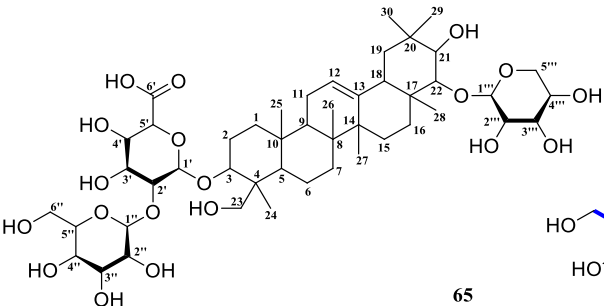
Table 5.5 Information of 22 compounds from DNP
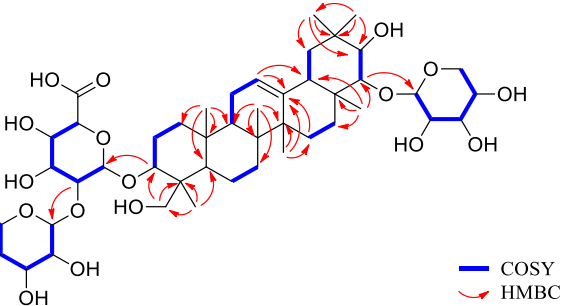
NO	Chemical name	Molecular weight	Molecular formula	Structure
1	Cimifoetiside V	944.4981	$C_{47}H_{76}O_{19}$	
2	Carnosifloside V	944.5345	$C_{48}H_{80}O_{18}$	
3	Carnosifloside VI	944.5345	$C_{48}H_{80}O_{18}$	
4	Scandenoside R6	944.5345	$C_{48}H_{80}O_{18}$	
5	Scandenoside R7	944.5345	$C_{48}H_{80}O_{18}$	
6	Quinquenoside L1	944.5345	$C_{48}H_{80}O_{18}$	
7	Gymnemaside II	944.5345	$C_{48}H_{80}O_{18}$	
8	ODB66-T	944.4981	$C_{47}H_{76}O_{19}$	
9	BFY75-H	944.5345	$C_{48}H_{80}O_{18}$	
10	Ilexoside I	944.4981	$C_{47}H_{76}O_{19}$	
11	PKN24-V	944.4981	$C_{47}H_{76}O_{19}$	
12	Subproside II	944.4981	$C_{47}H_{76}O_{19}$	

13	Clinopodoside H	944.5345	$C_{48}H_{80}O_{18}$	
14	Lupinoside PA2	944.4981	$C_{47}H_{76}O_{19}$	
15	**Kudzusaponin SA2**	**944.4981**	**$C_{47}H_{76}O_{19}$**	
16	Impatienoside B	944.4981	$C_{47}H_{76}O_{19}$	
17	Gymnemoside W1	944.5345	$C_{48}H_{80}O_{18}$	
18	Phaseoluside A	944.5345	$C_{48}H_{80}O_{18}$	
19	Derrissaponin	944.5345	$C_{48}H_{80}O_{18}$	
20	Solidagosaponin III	944.4981	$C_{47}H_{76}O_{19}$	
21	Mimusopside A	944.4981	$C_{47}H_{76}O_{19}$	
22	Ilexoside XXIII	944.4981	$C_{47}H_{76}O_{19}$	

Figure 5.13 Structures and observed COSY and key HMBC correlations of kudzusaponin (**65**).

Table 5.6 ^{1}H and ^{13}C (600 MHz) NMR data of **65** in CD$_3$OD

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1	38.3, CH$_2$	1.62 m	1', 2, 2'	5
		1.03 m	1, 2, 2'	
2	25.6, CH$_2$	2.11 m	1, 1', 2', 3	
		1.81 m	1, 1', 2, 3	10
3	90.5, CH	3.42 d (3.4)	2, 2'	23, 24, Glc-1
4	43.2, C			
5	55.9, CH	0.95 m	6, 6'	4, 6, 7, 10, 23, 25
6	18.0, CH$_2$	1.63 m	5, 6', 7, 7'	5, 8, 10
		1.38 m	5, 6, 7, 7'	5
7	32.4, CH$_2$	1.56 m	6, 6', 7'	8
		1.34 m	6, 6', 7	8
8	39.7, C			
9	47.5, CH	1.59 d (8.3)	11	1, 10, 11, 25, 26
10	36.0, C			
11	23.5, CH$_2$	1.89 m	9, 12	9
12	122.1, CH	5.20 t (3.0)	11	9, 13, 14, 18
13	143.6, C			
14	41.3, C			
15	25.9, CH$_2$	1.88 m	15', 16, 16'	13
		1.00 m	15, 16, 16'	13
16	27.1, CH$_2$	1.88 m	15, 15', 16'	18
		1.02 m	15, 15', 16	18
17	38.6, C			
18	43.8, CH	2.17 dd (3.5, 13.8)	19, 19'	12, 13
19	46.5, CH$_2$	1.91 m	18, 19'	18, 20, 29, 30
		1.41 m	18, 19	18, 20, 29, 30
20	35.7, C			
21	75.4, CH	3.52 m	22	29, 30
22	91.9, CH	3.33 d (3.2)	21	21, Ara-1
23	62.9, CH$_2$	4.12 d (11.5)	23'	3, 5
		3.22 d (11.5)	23	3, 4
24	21.6, CH$_3$	1.22 s		3, 4, 5, 23
25	14.8, CH$_3$	0.89 s		1. 5, 9, 10
26	15.9, CH$_3$	0.98 s		7, 8, 9, 14
27	25.5, CH$_3$	1.17 s		8, 13, 14, 15,

28	21.7, CH$_3$	0.99 s		16, 17, 18, 22
29	19.7, CH$_3$	1.02 s		20, 21, 30
30	29.9, CH$_3$	0.93 s		19, 20, 21, 29
Glc-1	103.5, CH	4.48 d (7.2)	Glc-2	3
Glc-2	79.4, CH	3.60 m	Glc-1, Glc-2	Glc-3, Gal-1
Glc-3	77.0, CH	3.61 m	Glc-2, Glc-3	Glc-2
Glc-4	75.2, CH	3.59 m	Glc-3, Glc-5	Glc-6
Glc-5	71.8, CH	3.48 m	Glc-4	
Glc-6	174.4, C			
Gal-1	103.3, CH	4.69 d (7.9)	Gal-2	Glc-2
Gal-2	72.0, CH	3.54 m	Gal-1, Gal-3	Gal-1, Gal-3
Gal-3	73.7, CH	3.43 m	Gal-2, Gal-4	Gal-2
Gal-4	69.8, CH	3.76 m	Gal-3,	Gal-2, Gal-3
Gal-5	75.4, CH	3.52 m	Gal-6, Gal-6'	
Gal-6	61.3, CH$_2$	3.79 dd (6.7, 11.8)	Gal-5, Gal-6'	Gal-4, Gal-6
		3.68 dd (4.8, 11.9)	Gal-5, Gal-6	Gal-4, Gal-6
Ara-1	107.4, CH	4.24 d (7.0)	Ara-2	22, Ara-3
Ara-2	72.4, CH	3.61 m	Ara-1, Ara-3	Ara-1, Ara-3
Ara-3	73.9, CH	3.48 m	Ara-2, Ara-4	Ara-1, Ara-4
Ara-4	68.7, CH	3.73 m	Ara-3, Ara-5, Ara-5'	Ara-2
Ara-5	66.2, CH$_2$	3.84 dd (2.6, 12.6)	Ara-4, Ara-5'	Ara-3, Ara-4
		3.48 m	Ara-4, Ara-5	Ara-3, Ara-4

In the serial dilution experiment, the MICs of borrelidin (**64**) and kudzusaponin SA2 (**65**) against BCG were determined to be <1.5 and 106 μM. When compared with the positive control isoniazid, **64** exhibited almost equal MIC on BCG, indicating that this compound could be an alternative drug candidate for treating tuberculosis. Furthermore, the effects of **64** on *Staphylococcus aureus* (SA) and *Bacillus subtilis* (BS) were also tested. The inhibitory activities were observed at 204 and 51 μM, respectively.

When comparing proton NMR spectra of active fractions and pure compounds, almost all of the signals have been aligned with those in the spectra of compounds **64** and **65**, except for signals larger than 7 ppm (Figure 5.14). From their coupling types, signals larger than 7 don't belong to analogs of either compound, indicating another class of metabolites may be produced by LS120167.
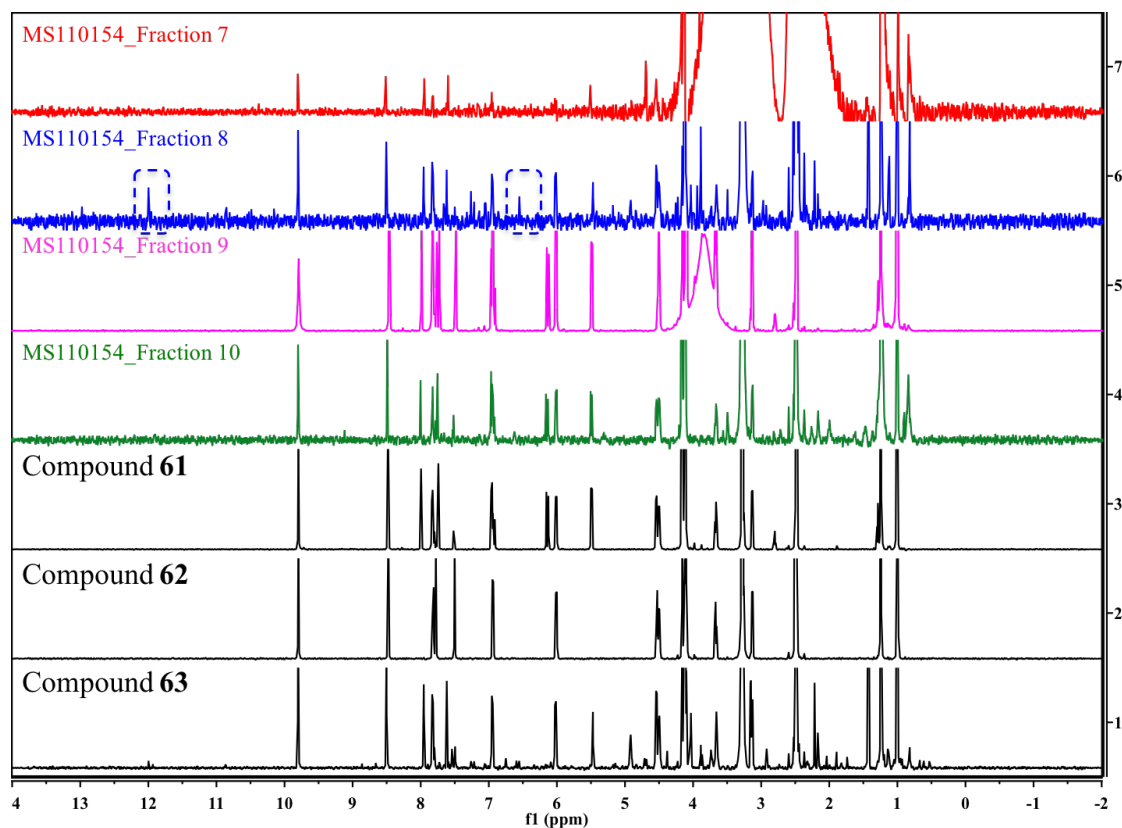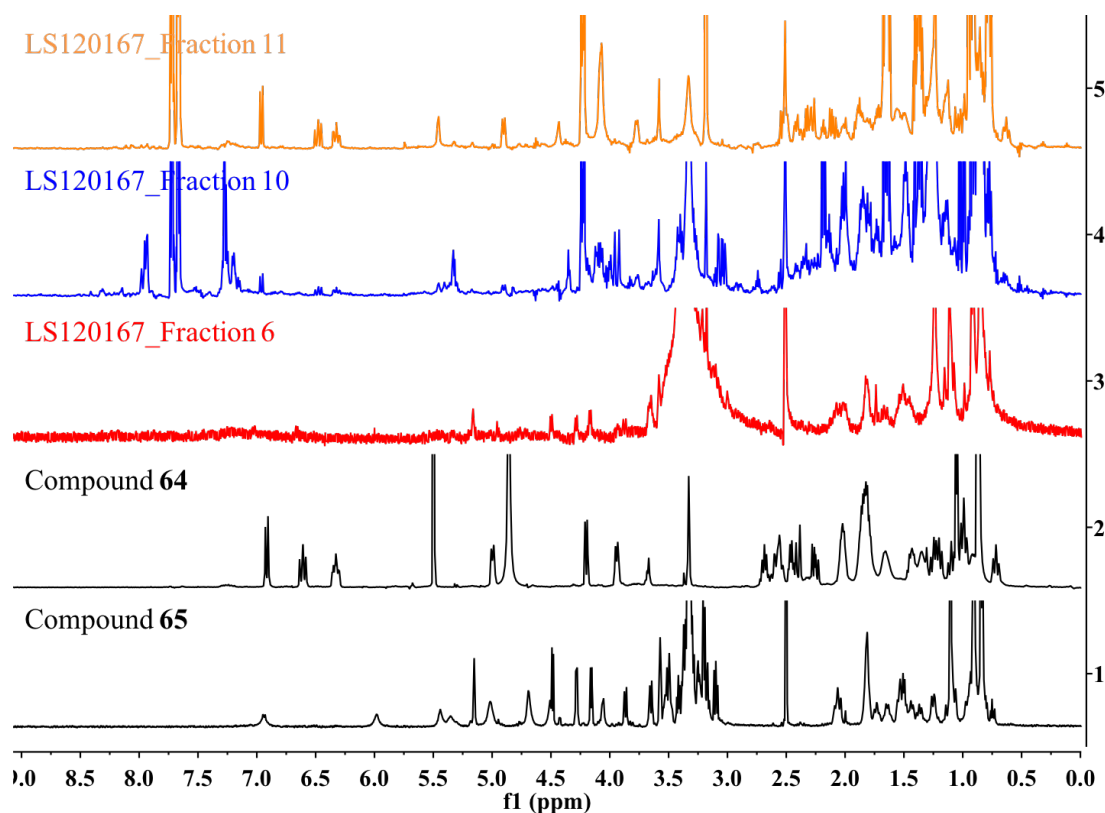
Figure 5.14 Comparison of NMR fingerprint between active fractions and compound of
LS120167_AM2.

5.4.3 Small-scale isolation of ES130159_M001

Endophyte ES130159 was isolated from TCM *Cicuta virosa,* which was collected from
Sichuan Province, China. It was identified as *Streptomyce* based on the morphology and
16S ribosomal DNA (rDNA) phylogenetic analysis (Figure 5.15).
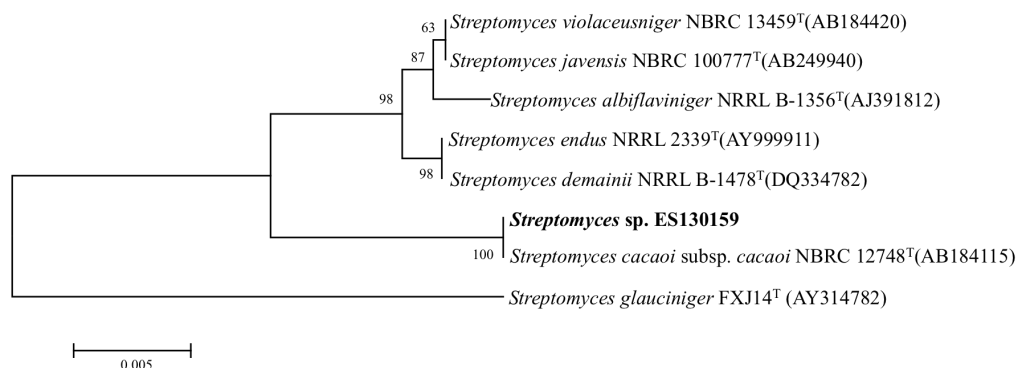
Figure 5.15 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of ES130159 and
related species constructed by MEGA.

Fraction 3 and 4 were found to be active in anti-BCG screening. ^{1}H NMR fingerprint of the active fractions revealed coupled NMR signals in the unsaturated region as well as several methyl and methylene groups (Figure 5.16).
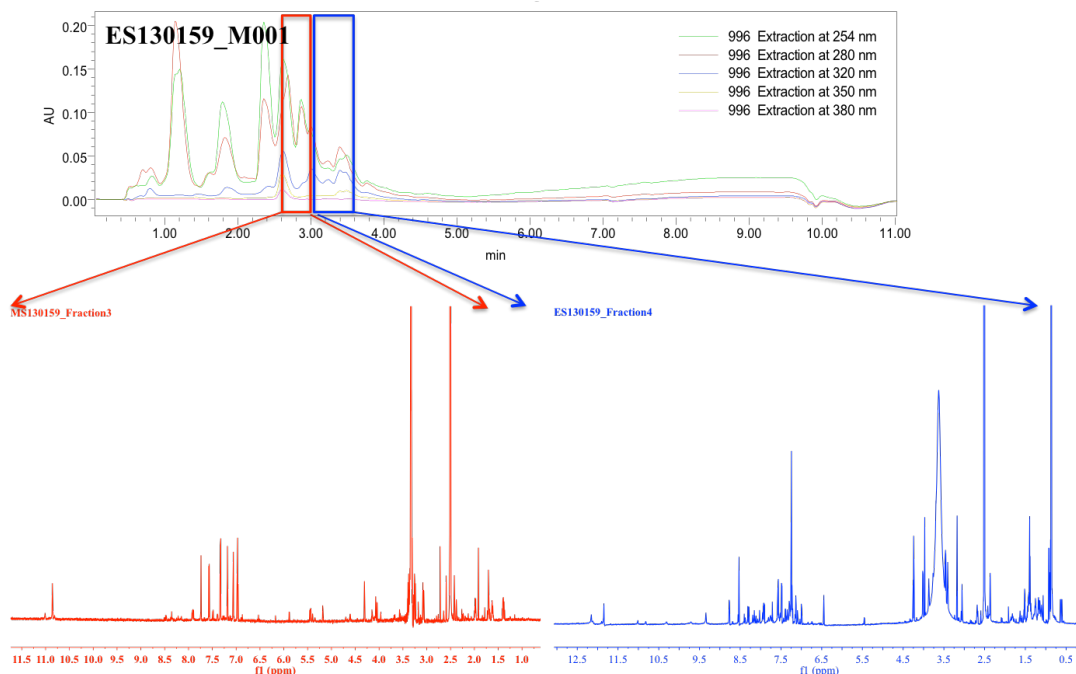
Figure 5.16 HPLC trace and NMR fingerprint of ES130159_M001.

Compound **66** (Figure 5.17) was obtained as a white powder and its molecular weight was determined to be 295 based on LC-MS. The NMR data displayed signals for 2 aromatic protons (δ_H 8.22; δ_C 150.7; δ_H 8.38; δ_C 140.0), a sugar moiety with 4 methines (δ_H 5.92; δ_C 88.2; δ_H 4.58; δ_C 74.0; δ_H 4.15; δ_C 71.0; δ_H 3.96; δ_C 86.2), one methylene (δ_H 3.59, 3.68; δ_C 62.2), three hydroxyl groups (δ_H 5.46; 5.20; 5.37), and two methyl groups (δ_H 3.17; δ_C 49.1; δ_H 3.18; δ_C 49.1). The positions of two methyls and sugar moiety were determined by HMBC correlations from H-8, H-9 to C-7 and H-5 to C1′ (Table 5.7). Absolute configuration of compound **66** was determined by comparison the [α] $_D$ value in methanol with literature.[148]

Figure 5.17 Structures and observed COSY and key HMBC correlations of **66**.

Table 5.7 ^{1}H and ^{13}C (600 MHz) NMR data of **66** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
2	150.7 CH	8.22 s		3a, 7
3a	150.0 C			
5	140.0 CH	8.38 s		3a, 6a, 1'
6a	155.3 C			
8	119.5 C			
7	120.6 C			
8	49.1 CH$_3$	3.17 s		7
9	49.1 CH$_3$	3.18 s		7
1'	88.2 CH	5.92 d (6.0)	2'	3', 4'
2'	74.0 CH	4.58 dd (5.76, 10.93)	1', 3' 3'-OH	4'
3'	71.0 CH	4.15 dd (4.53, 8.42)	2', 4', 3'-OH	1', 5'
4'	86.2 CH	3.96 dd (3.40, 6.79)	3', 5'	4'
5'	62.2 CH$_2$	3.59 m	4', 5'-OH	3'
		3.68 m		3'
2'-OH		5.46 d (5.18)		
3'-OH		5.20 d (4.74)		
5'-OH		5.37 dd (4.79, 7.01)		

67 was obtained as a white powder and its molecular weight 283 determined by LC-MS. ^{1}H NMR revealed two exchangeable protons at δ_H 7.74 and 10.85 ppm, four correlated aromatic protons at δ_H 7.57 (d, J=7.94 Hz), 7.33 (td, J= 0.83, 8.07 Hz), 7.06 (dt, J=1.10, 7.52 Hz) and 6.97 (dt, J=1.00, 7.42), suggesting a 1,6-disubstituted phenyl ring moiety, two down-field methine at δ_H 4.31 (t, J=5.45 Hz) and 7.07 (dt, J=1.00, 8.38) and four methelenes at δ_H 1.62 ~ 3.28 ppm (Table 5.8). 13C NMR data showed the existence of two carbonyl carbons at δ_C 165.9 and 169.4 ppm. Full assignments of the NMR data of **67** were achieved by means of 2D NMR experiments, and the key COSY and HMBC correlations are shown in Figure 5.18. HMBC correlations from methylene protons at 3.38 and 3.26 ppm to the carbonyl carbon at δ_C 165.9 ppm indicated the connection of C-15 and C-16 via nitrogen. The location of the other carbonyl at 169.4 ppm was confirmed by HMBC correlations from H-10, H-12 and H-13'. Comparison with published data confirmed the structure of compound **67** as a known compound brevianamide F.[149]

Table 5.8 ^{1}H and ^{13}C (600 MHz) NMR data of **67** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1		10.85 s		2, 3, 3a, 7a
2	124.3 CH	7.18 d (2.40)		3, 3a, 7a, 8
3	109.8 C			
3a	127.8 C			
4	119.1 CH	7.57 d (7.94)	5	3, 5, 3a, 6a
5	118.7 CH	6.97 dt (1.00, 7.42)	4, 6	3a, 6, 7
6	121.3 CH	7.06 dt (1.10, 7.52)	5, 7	2, 4, 7a
7	111.7 CH	7.33 td (0.83, 8.07)	6	3a, 5
7a	136.4 C			
8	26.3 CH$_2$	3.24 m	8', 9	2, 3, 3a, 9, 16
		3.08 dd (5.92, 15.72)	8, 9	2, 3, 3a, 9, 16
9	55.6 CH	4.31 t (5.45)	8, 8'	3, 16
10		7.74 s		8, 9, 11, 12, 16
11	169.4 C			
12	58.9 CH	4.07 dt (1.00, 8.38)	13, 13'	11, 13
13	28.1 CH$_2$	1.98 m	12, 13', 14, 14'	15
		1.39 m	12, 13, 14, 14'	11, 12, 14, 15
14	22.3 CH$_2$	1.70 m	13, 13', 14', 15, 15'	13, 15
		1.62 m	13, 13', 14, 15, 15'	12
15	45.1 CH$_2$	3.38 m	14, 14',15'	12, 13, 14
		3.26 m	14, 14',15	13, 14
16	165.9 C			
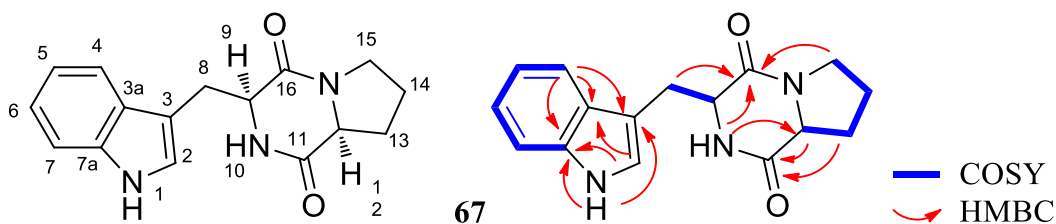
Figure 5.18 Structures and observed COSY and key HMBC correlations of **67**.

Compound **68**, obtained as yellow powder, was assigned the molecular weight as 279 by LC-MS and with the consideration of NMR data, its molecular formula was determined

as $C_{16}H_9NO_4$. The ^1H NMR spectrum revealed five aromatic protons, including one aromatic singlet (δ_H 7.24; δ_C 112.1) and four coupled aromatic protons (δ_H 8.53; δ_C 128.3; δ_H 7.57; δ_C 126.6; δ_H 7.48; δ_C 126.9; δ_H 8.78; δ_C 128.1), two exchangeable singlets of a H-bonded phenolic hydroxyl (δ_H 12.09), and an amidocarbonyl proton at (δ_H 9.29), and a methylenedioxy group (δ_H 6.47; δ_C 101.8). Five aromatic protons of the spectrum closely resembled those of other 7,8,9,10,11-unsubstituted dioxoaporphine (Table 5.9). The singlet at δ 8.53 was ascribed to H-7 while H-8, H-9, H-10 and H-11 appeared as an ABCX coupling system. In agreement with the spectra of other C- 11-unsubstituted aporphines, dehydro- and oxoaporphines, the most downfield doublet (δ_H 8.78) was characteristic of H-11, the protons (δ_H 7.48, 7.57) were assigned to H-10, 9, and another proton (δ_H 8.53) was due to H-8. These data indicated that **68** was a trisubstituted aristolactam derivative. Thus, the structure of **68** was determined to be a known natural product pressalanine A (Figure 5.19).

Compound **69**, obtained as yellow powder, was determined the molecular weight as 307 by LC-MS and with the consideration of NMR data, its molecular formula was determined as $C_{17}H_9NO_5$ (Figure 5.19). The NMR spectra of **69** showed very similar pattern with those of **68**, indicating that **69** was a also a trisubstituted dioxoaporphine derivative. By comparing of the NMR data with **68** and the literature[150], the structure of 69 was determined to be a known compound pressalanine B. The presence of the additional carbonyl group was also confirmed by HMBC correlations from H-7 (δ_H 9.36) to C-5 (δ_C 176.8) and C-6 (δ_C 157.0) (Table 5.9).
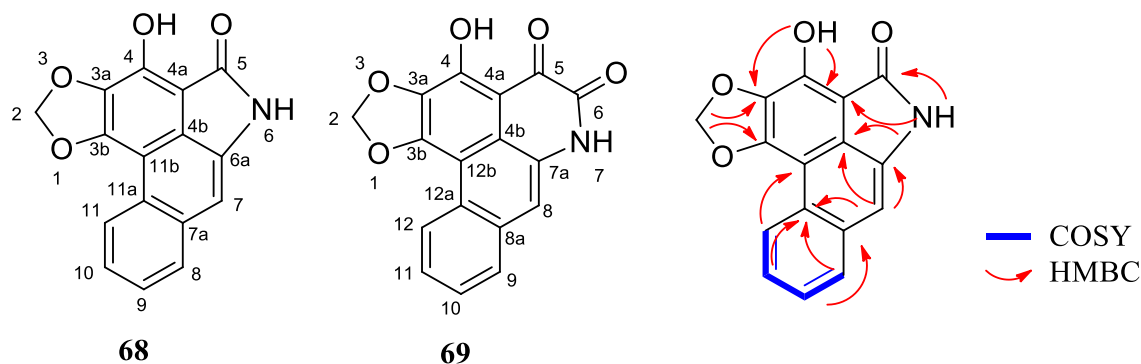
Figure 5.19 Structures and observed COSY and key HMBC correlations of pressalanine A (**68**) and B (**69**)

Table 5.9 ^{1}H and ^{13}C (600 MHz) NMR data of **68** and **69** in DMSO-d_6

	68					**69**			
Position	δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC	Position	δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC
2	101.8 CH$_2$	6.47 s		3a, 3b	2	101.6 CH$_2$	6.42 s		3a, 3b
3a	132.7 C				3a	133.1 C			
3b	149.2 C				3b	151.6 C			
4	170.7 C				4	173.1 C			
4a	113.8 C				4a	102.5 C			
4b	125.0 C				4b	120.9 C			
5	169.6 C				5	176.8 C			
6		9.29 s		4a, 4b, 5	6	157.0 C			
6a	136.3 C				7		9.36 s		4b, 5, 6, 8
7	112.1 CH	7.24 s		4b, 6a, 11a	7a	131.8 C			
7a	135.5 C				8	112.1 CH	7.22 s		4b, 7a, 12a
8	128.3 CH	8.53 dd (1.64, 8.12)	9	7, 9, 10, 11a	8a	133.5 C			
9	126.6 CH	7.57 t (8.14)	8, 10	7a, 8, 10, 11	9	128.3 CH	8.28 dd (1.49, 8.04)	9	8, 10, 11, 12a
10	126.9 CH	7.48 t (7.96)	9,11	8, 11, 11a	10	126.7 CH	7.70 t (8.24)	8, 10	8a, 9, 12
11	128.1 CH	8.78 dd (1.51, 7.77)	10	7a, 11a, 11b	11	126.6 CH	7.55 t (7.86)	9,11	9, 12, 12a
11a	127.0 C				12	128.4 CH	8.30 dd (1.49, 7.84)	10	8a, 12a, 12b
11b	116.9 C				12a	127.5 C			
4-OH		12.09 s		3a, 4a	12b	117.3 C			
					4-OH		12.10 s		3a, 4a

Compound **70** was obtained as white powder. The molecular weight was determined as 210 based on LC-MS result and its molecular formula was identified as $C_{12}H_{18}O_3$ with the consideration of NMR data (Table 5.10). The ^{1}H-NMR spectrum of **70** displayed two methyl groups (δ_H 1.38; δ_C 16.1; δ_H 0.87; δ_C 18.9), one olefinic signal (δ_H 5.86; δ_C 124.0), and two oxygenated CH (δ_H 4.19; δ_C 74.5; δ_H 4.59; δ_C 71.7). In the HMBC spectrum, correlations of H-6 with C-5, C-7, C-8, C-10 and C-11, and of H-9) with C-1, C-5 and C-7, confirmed the presence of an a,b,a',b'-unsaturated ketone moiety. Correlations of the H-atom of a OH group (δ_H 4.10) with C-6, C-7 and C-8 confirmed the location of OH at C-7, correlations of another OH group (δ_H 4.00) with C-1, C-2 and C-10 indicated confirmed the location of OH at C-1. The above discussed analysis of the 1D- and 2D-

NMR data further identified the structure of compound **70** (Figure 5.20). Absolute configuration of compound **70** was determined by comparison the $[\alpha]_D$ value in methanol with literature.[151]
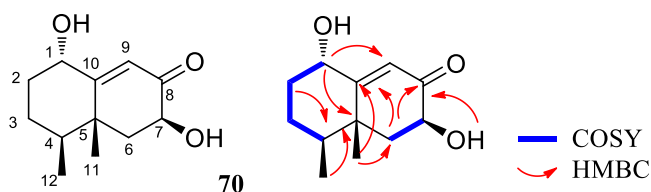
Figure 5.20 Structures and observed COSY and key HMBC correlations of compound **70**.

Table 5.10 ^{1}H and ^{13}C (600 MHz) NMR data of **70** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1	74.5 CH	4.19 t (8.14)	2, 1-OH	3, 5, 9
2	35.2 CH$_2$	2.12 m	1, 2', 3, 3'	4, 10
		1.90 m	1, 2, 3, 3'	3, 4, 10
3	25.5 CH$_2$	1.87 m	2, 2', 3', 4	1, 5, 12
		1.36 m	2, 2', 3', 4	1, 12
4	43.4 CH	1.78 m	3, 3', 12	2, 10, 12
5	45.1 C			
6	43.2 CH2	2.13 m	6', 7	5, 7, 8, 10, 11
		1.93 m	6, 7	5, 7, 8, 10, 11
7	71.7 CH	4.59 dd (14.1, 5.4)	6, 6', 7-OH	5, 6, 9
8	194.8 C			
9	124.0 CH	5.86 s		1, 5, 7
10	169.8 C			
11	16.1 CH$_3$	1.38 s		4, 5, 6, 10
12	18.9 CH$_3$	0.87 d (6.24)	4	3, 4, 5
1-OH		4.00 d (5.16)	1	1, 2, 10
7-OH		4.10 d (5.21)	7	6, 7, 8

Compound **71** was isolated as white powder. The molecular weight was determined as 161 based on LC-MS result and molecular formula was assigned as $C_9H_7NO_2$ with the combinational analysis of NMR data (Figure 5.21). The 1D NMR spectra of **71** showed resonances for five aromatic protons (δ_H 7.99; δ_C 131.7; δ_H 7.46; δ_C 112.3; δ_H 7.18; δ_C 122.3; δ_H 7.15; δ_C 121.8; δ_H 8.00; δ_C 120.7) and two exchangeable protons (δ_H 11.91; 11.80). Together with the analysis of 2D NMR, compound **71** contained a typical benzopyrrole skeleton. The location of carboxyl group was determined by HMBC

correlations from H-3 to C-8 and 8-OH to C-2 (Table 5.11).
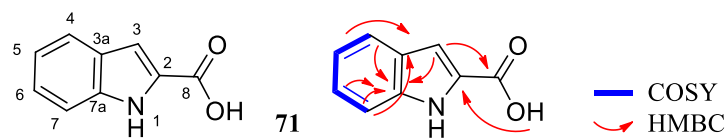
Figure 5.21 Structures and observed COSY and key HMBC correlations of **71**.

Table 5.11 ^{1}H and ^{13}C (600 MHz) NMR data of **71** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1		11.91 brs		
2	126.2 C	11.9 s		
3	131.7 CH	7.99 d (2.91)		4, 3a, 7a, 8
3a	130.7 C			
4	112.3 CH	7.46 dt (0.92, 7.84)	5	3, 6, 7a
5	122.3 CH	7.18 td (1.38, 7.08)	4, 6	3a, 7
6	121.8 CH	7.15 td (1.11, 7.06)	5, 7	4, 7a
7	120.7 CH	8.00 d (7.70)	5	3a, 5
7a	132.0 C			
8	160.3 C			
8-OH		11.8		2

Compound **72** was obtained as white powder. Its molecular weight was determined as 176 based on LC-MS data and molecular formula was assigned as $C_9H_8N_2O_2$ with the analysis of NMR data (figure 5.22). The 1D NMR spectra of **72** showed signals for three aromatic protons (δ_H 7.90; δ_C 127.2; δ_H 7.48; δ_C 135.5; δ_H 7.92; δ_C 124.6) one methyl group (δ_H 2.72; δ_C 19.2) and two exchangeable protons (δ_H 7.91; 8.34). Together with the analysis of 2D NMR, compound **72** displayed a typical quinazoline skeleton. The substitute positions of methyl group and two hydroxyl groups were determined by HMBC correlations from H-9 to C-5, C-6 and C-4a, and 2-OH to C-2, and 4-OH to C-3 and C-4a (Table 5.12).
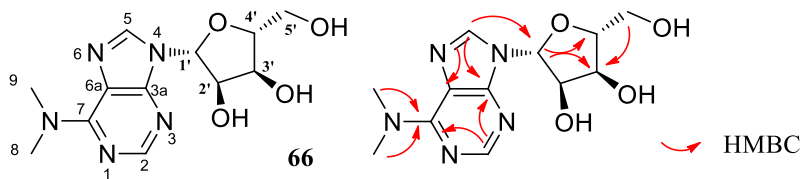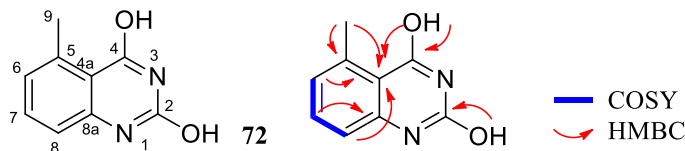

Figure 5.22 Structures and observed COSY and key HMBC correlations of **72**.

Table 5.12 ^{1}H and ^{13}C (600 MHz) NMR data of **72** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
2	187.2			
4	177.1			
4a	110.1			
5	126.5			
6	127.2	7.90 dd (1.01, 8.05)	7	4z, 8, 9
7	135.5	7.48 t (7.77)	6, 8	5, 8a
8	124.6	7.92 dd (1.07, 7.71)	7	4a, 6, 8a
8a	151.0			
9	19.2	2.72 s		4a, 5, 6
2-OH		7.91 brs		2
4-OH		8.34 brs		4, 4a

Seven compounds isolated from ES130159 were assayed against various microbial pathogens (Table 5.13). Only compounds **67**, **68** and **69** exhibited weak anti-BCG and anti-SA activities. Compounds **67** and **68** also displayed weak anti-MRSA activities with MICs of 28.3 and 27.9 μM, respectively.

Table 5.13 Antimicrobial activities of compounds

MIC (μM)	Organism (Strain)				
	BCG	SA	MRSA	BS	PA
66	NA	NA	NA	NA	NA
67	14.15	28.3	28.3	NA	NA
68	27.9	27.9	27.9	NA	NA
69	30.7	30.7	NA	NA	NA
70	NA	NA	NA	NA	NA
71	NA	NA	NA	NA	NA
72	NA	NA	NA	NA	NA
Control	0.37[a]	0.7[b]	0.7[b]	0.35[b]	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

By comparison of NRM fingerprint of between active fractions and identified compounds, almost all signals in active fractions were isolated (Figure 5.23).
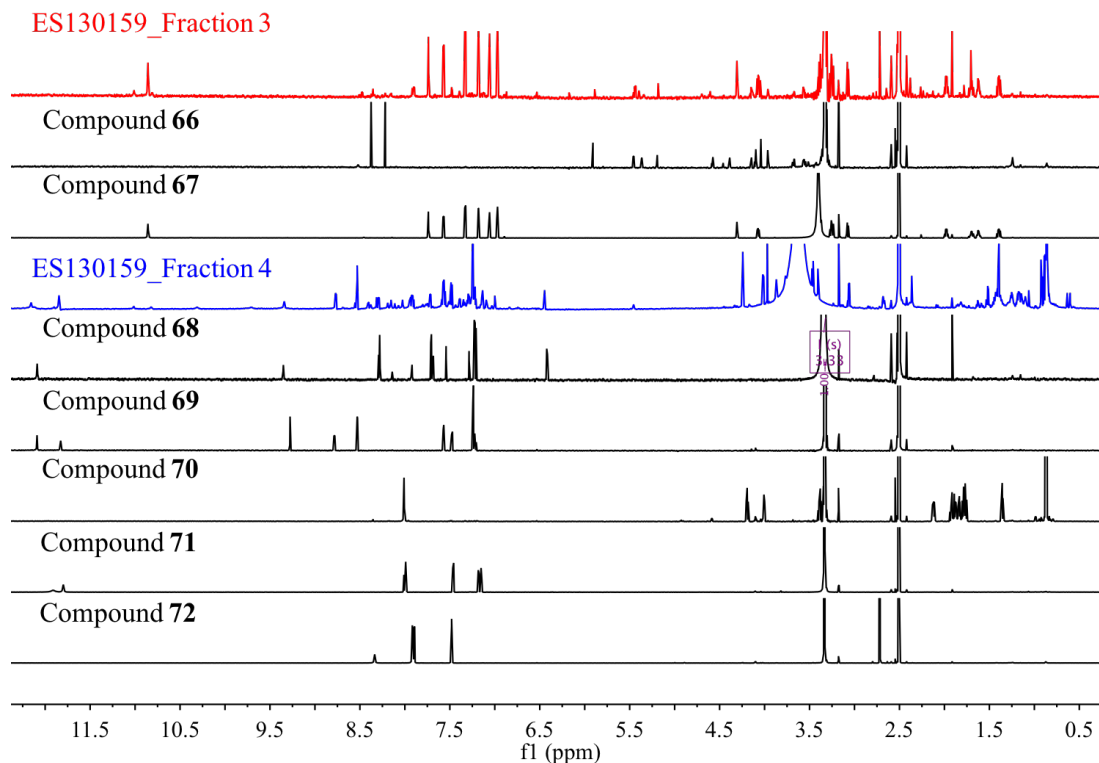
Figure 5.23 Comparison of NMR fingerprint between active fractions and compound of ES130159_M001.

5.4.4 Small-scale isolation of MS110105_MPG

Marine actinomycete MS110105 was isolated from a sediment sample collected from the South China Sea. This isolate was identified as a *Streptomyces* sp. based on cultural, physiological, morphological characteristics and 16s rRNA gene analysis (Figure 5.24).
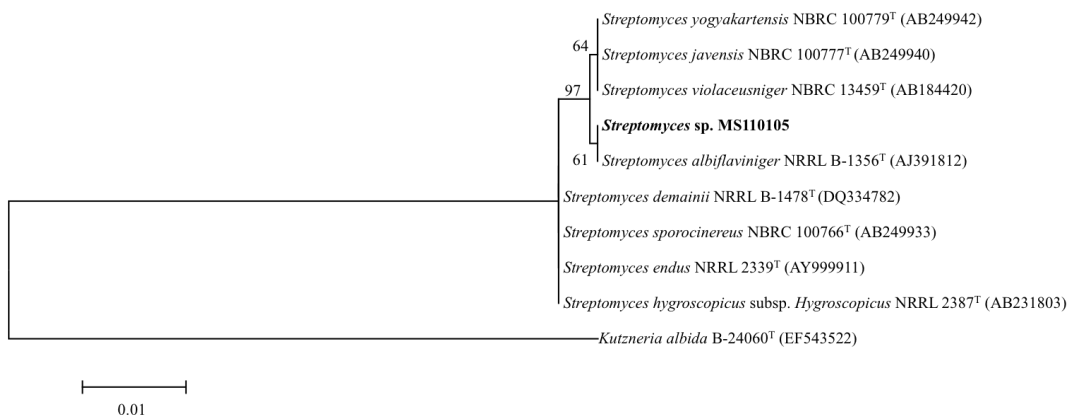
Figure 5.24 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110105 and related species constructed by MEGA.

While fraction 8 and 9 of MS110105_MPG didn't show enough information in the small-

scale ^{1}H NMR fingerprint because of the relatively low yield, fraction 10 on the other hand, contained signals in both the aliphatic and the aromatic regions (Figure 5.25).
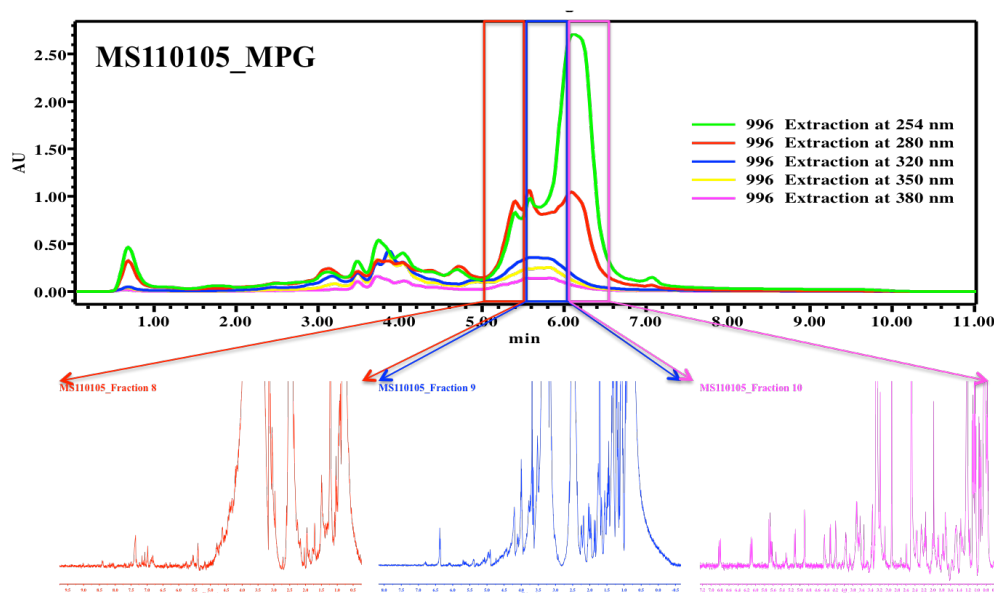
Figure 5.25 HPLC trace and NMR fingerprint of MS110105_MPG.

HPLC fractionation of the active fractions gave pure elaiophylin (**73**) and elaiophylin G (**74**) (Figure 5.26).

73: R=CH$_3$
74: R=H
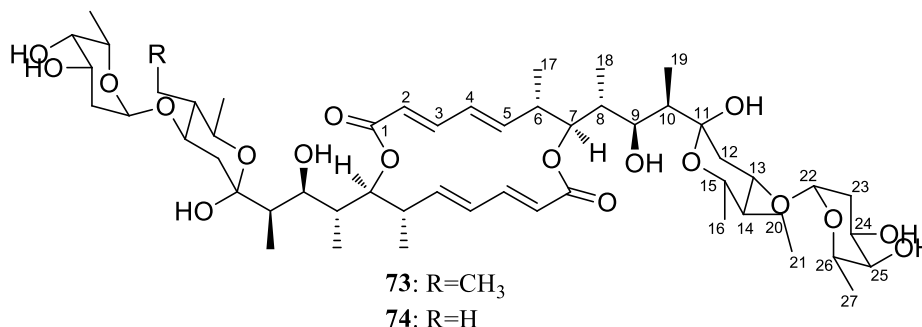
Figure 5.26 Structures of elaiophylin (**73**) and elaiophylin G (**74**).

Elaiophylin (**73**) was obtained as colorless powder that gave a (M +H)$^+$ peak at m/z 1025 in the LC-MS appropriate for a molecular formula of C$_{54}$H$_{88}$O$_{18}$. It gave only half the proton and carbon signals respectively in the ^{1}H and ^{13}C-NMR spectra in CDCl$_3$ solution (Table 5.14), revealing the presence of symmetry in its molecular structure. Detailed analyses of the ^{1}H and ^{13}C-NMR spectra and 2D-NMR (COSY, HSQC, HMBC) techniques led us consider **73** to be elaiophylin, a C2-symmetric macrodiolide with a 16-membered dilactone ring. Eventually, comparing with the NMR data of **73** and elaiophylin reported before confirmed that **73** is elaiophylin.

Table 5.14 ^{1}H and ^{13}C (600 MHz) NMR data of **73** and **74** in CDCl$_3$

Position	73		74	
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)
1, 1'	169.9 C		170.0 C	
2, 2'	121.1 CH	5.69 d (15.5)	121.0 CH	5.70 d (16.0)
3, 3'	145.1 CH	6.98 dd (15.5, 11.0)	145.0 CH	6.99 dd (15.8, 11.5)
4, 4'	132.0 CH	6.12 dd (15.1, 11.2)	132.0 CH	6.13 dd (15.1, 11.5)
5, 5'	144.4 CH	5.63 dd (15.1, 10.2)	144.3 CH	5.64 dd (15.0, 10.0)
6, 6'	40.9 CH	2.54 m	40.8 CH	2.55 m
7, 7'	77.8 CH	4.75 m	77.9 CH	4.74 m
8, 8'	36.0 CH	1.96 m	35.9 CH	1.96 m
9, 9'	70.7 CH	4.11 m	70.6 CH	4.11 m
10, 10'	41.8 CH	1.71 m	41.6, 41.5 CH	1.72 m
11, 11'	99.1 C		99.1, 99.0 C	
12, 12'	39.0 CH$_2$	2.37 dd (4.5, 11.2)	38.8, 38.5 CH$_2$	2.39 dd (4.5, 11.2)
		1.03 m		1.03 m
13, 13'	70.5 CH	3.96 m	70.0, 73.3 CH	3.97, 3.66 m
14, 14'	48.6 CH	1.18 m	48.3, 43.5 CH	1.21, 1.29 m
15, 15'	66.7 CH	3.90 m	66.5, 69.3 CH	3.92, 3.71 m
16, 16'	19.2 CH$_3$	1.10 d (6.4)	19.4, 19.1 CH$_3$	1.10, 1.11 d (6.4)
17, 17'	15.0 CH$_3$	1.04 d (6.7)	14.9 CH$_3$	1.04 d (7.1)
18, 18'	8.8 CH$_3$	0.81 d (7.1)	8.8 CH$_3$	0.82 d (7.1)
19, 19'	7.1 CH$_3$	1.00 d (7.4)	7.1 CH$_3$	1.01 d (7.0)
20	19.5 CH$_2$	1.61 m	24.9 CH$_2$	1.63 m
		1.44 m		1.42 m
20'	19.5 CH$_2$	1.61 m	13.4 CH$_3$	0.92 d (6.5)
		1.44 m	--	--
21	9.3 CH$_3$	0.86 t (7.6)	9.0 CH$_3$	0.85 t (6.5)
21'	9.3 CH$_3$	0.86 t (7.6)	--	--
22, 22'	93.4 CH	5.05 m	93.1, 92.9 CH	5.06 m
23, 23'	33.7 CH$_2$	1.90 m	33.5, 33.4 CH$_2$	1.80 m
24, 24'	66 CH	3.97 m	65.9, 65.8 CH	4.00 m
25, 25'	71.8 CH	3.61 m	71.4 CH	3.62 m
26, 26'	66.2 CH	3.99 m	66.0 CH	4.00 m
27, 27'	16.9 CH$_3$	1.24 d (6.7)	16.8, 16.7 CH$_3$	1.25 d (6.5)

Elaiophylin G (**74**) gave the UV absorptions closely similar to those of **73** and its

molecular formula, $C_{53}H_{86}O_{18}$, is CH_2 less than **73** ($C_{54}H_{88}O_{18}$), showing that **74** is a macrodiolide belonging to the same class. The ^{1}H-NMR spectrum of **73** in $CDCl_3$ solution showed signal patterns (both chemical shifts and splitting patterns) closely resembled those of 1, but it was characterized by the methyl signals (δ_H 0.92, d, J=6.5Hz, H-20′, δ_C 13.41, C-20′) in **74** instead of the ethyl signals (δ_H 0.86, t, J=7.6 Hz, H-21, δ_C 9.30, C-21; δ_H 1.61,m and 1.44, m, H-20, δ_C 19.53, C-20) in **73** and by the appearance of other signals as pairs having closely similar chemical shifts with the same splitting patterns. Each pair of ^{1}H signals appeared seemingly complex because of overlapping of split signals. Careful analyses of the ^{1}H NMR spectra by the combined use of 2D NMR experiments enabled us to dereplicate the structure of **74** and finally to identify as elaiophylin G.

The macrodiolide metabolites, elaiophylin (**73**) and elaiophylin G (**74**), were assayed toward various target organisms (Table 5.15). Both compounds showed antibacterial activity toward Gram- positive pathogens, including BCG, SA, MRSA and BS, but were inactive toward Gram- negative PA. Particularly, elaiophylin (**73**) showed significant antibacterial activities against the tested Gram- positive pathogens with the MIC values ranging from 0.8 to 1.5 μM. In contrast, elaiophylin G (**74**) was also found to have weak antibacterial activities (MIC 49 and 98 μM). The differences between the bacterial inhibitions of two compounds are possibly due to the different symmetrical parts in each structure, which resulted in different 3D structures.

Table 5.15 Antimicrobial activities of compounds

Organism (Strain)	Minimum Inhibitory Concentration (μM)		
	73	**74**	Control
Bacillus Calmette-Guérin (Pasteur 1173P2, BCG)	1.5	49	0.37[a]
Staphylococcus aureus (ATCC 6538)	<0.8	98	0.7[b]
Methicillin-resistant *S. aureus* (Clinical strain of Chaoyang hospital)	1.5	98	0.7[b]
Bacillus subtilis (ATCC 6633)	1.5	49	0.35[b]
Pseudomonas aeruginosa (PAO1)	NA*	NA	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

By comparison of the fingerprint of active fractions and identified compounds, several signals in aromatic area, which do not belong to the same class of compound **73** and **74**, were detected in fractions 9 and 10 (Figure 5.27). This may indicate a new class of

compounds from MS110105 in future investigation.
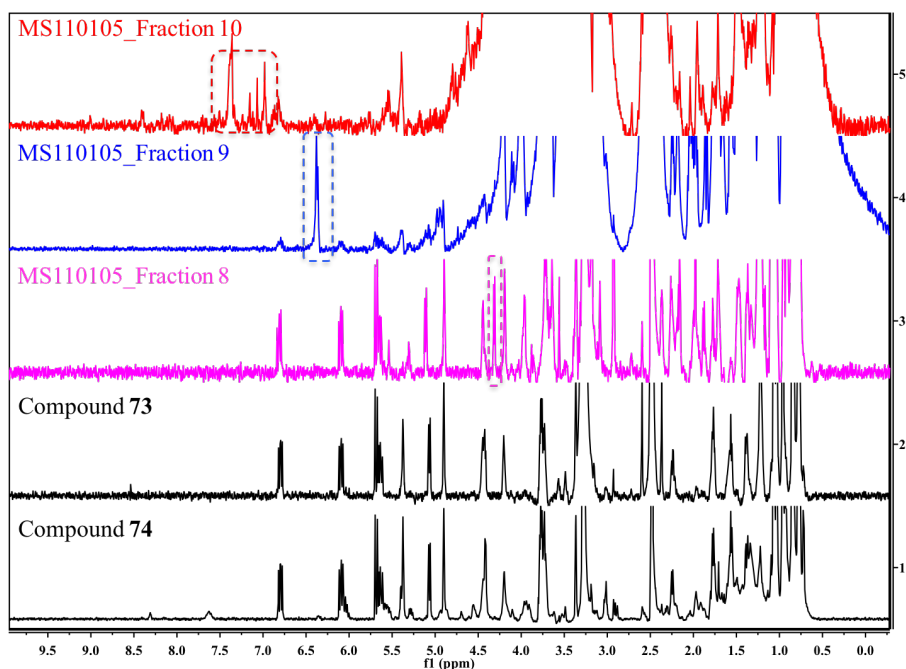
Figure 5.27 Comparison of NMR fingerprint between active fractions and compounds of MS110105_MPG.

5.4.5 Small-scale isolation of MS110109_MPG

Another marine actinomycete MS110109 was also isolated from a sediment sample collected from the South China Sea. It was identified as a *Streptomyces* sp. based on cultural, physiological, morphological characteristics and 16s rRNA gene analysis (Figure 5.28).
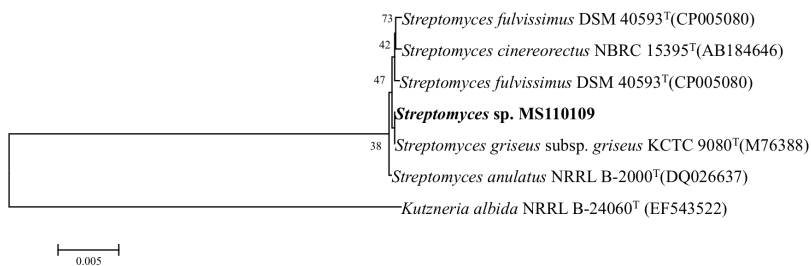
Figure 5.28 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110109 and related species constructed by MEGA.

The ^{1}H NMR fingerprint of fraction 7 of MS110109_MPG didn't show any significant signals, however, fractions 5, 6 and 8 displayed signals in the aromatic region and also some methyl signals, suggesting the presence of some small molecules in this extract (Figure 5.29).
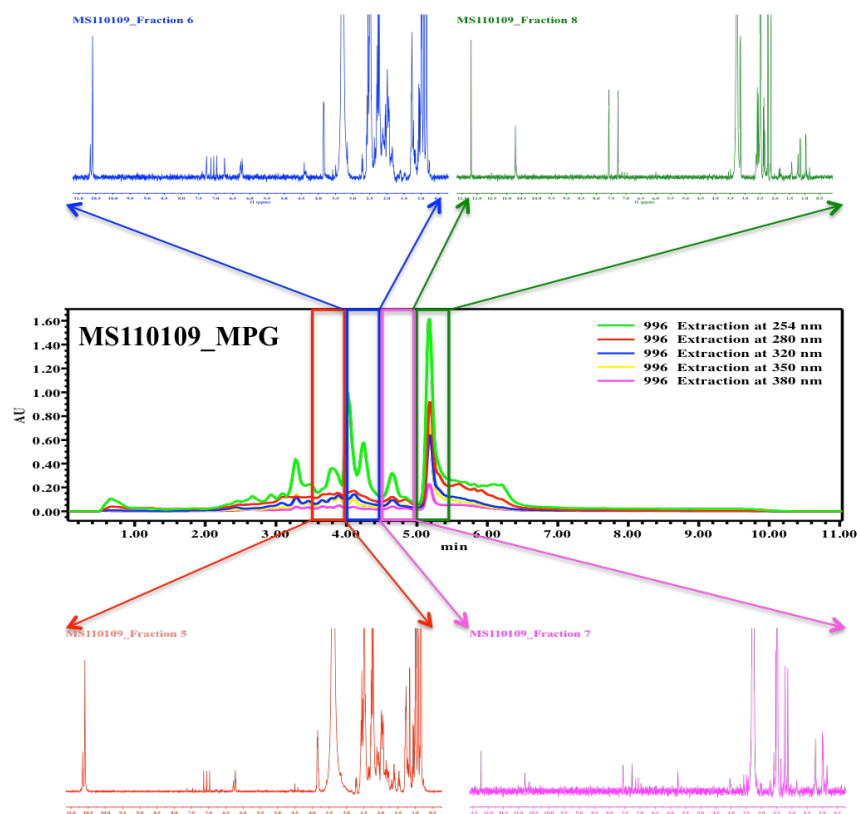
Figure 5.29 HPLC trace and NMR fingerprint of MS110109_MPG.

The active fractions were further separated by reversed-phase preparative HPLC using MeOH-H_2O as mobile phase to yield three known compounds actiphenol (**75**), nong-kang 101G (**76**) and cycloheximide (**77**) (Figure 5.30).

75: R=H
76: R=OH
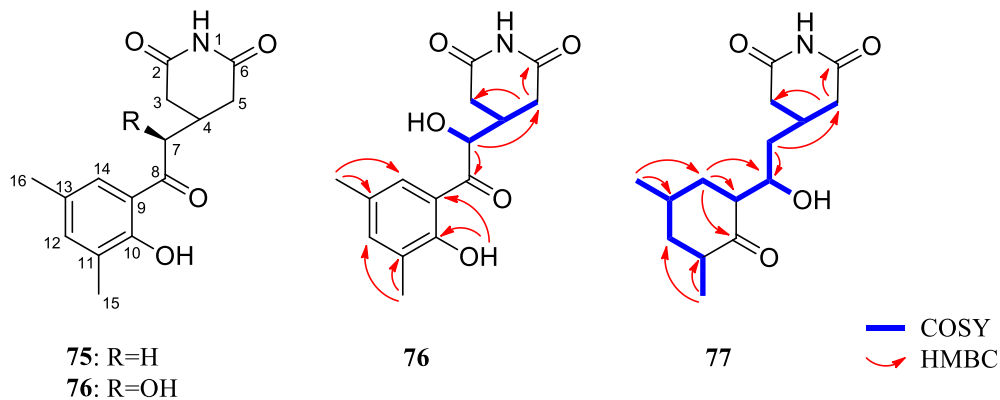
Figure 5.30 Structures and observed COSY and key HMBC correlations of actiphenol (**75**), nong-kang 101G (**76**) and cycloheximide (**77**).

The molecular formula of actiphenol (**75**) was determined to be $C_{15}H_{17}NO_4$ on the basis of LC-MS measurement. The ^{1}H NMR (DMSO-d_6) spectrum showed seventeen signals

due to two methyl, three methylene, three methine, one NH and one OH protons. The ^{13}C NMR (DMSO-d_6) spectrum exhibited fifteen signals consisting of two methyl, three methylene, one methine, six aromatic, and three carbonyl carbons. The correlation between protons and carbons was elucidated from the COSY spectra as shown in Table 5.9. The HMBC experiment showed the presence of an aromatic ring and a glutarimide ring (Figure 5.30). The proton signals of 2- and 4-CH$_3$ substituted to the aromatic ring were assigned by HMBC experiment (Table 5.16).

Nong-kang 101G (**76**), a colorless powder, had a molecular formula $C_{15}H_{17}NO_5$, as suggested from the LC-MS, indicating eight degrees of unsaturation. The ^1H and ^{13}C NMR spectra of **76** showed high analogy to those of actiphenol (**75**), except that **76** had a signal of an *O*-bearing methane C-atom at δ_C 70.7 (C-7) instead of a CH$_2$ signal. This suggested that **76** and **75** differ structurally only in the presence of an OH substituent in **76**, which was located at C7 based on the HMBC (Table 5.16).

Cycloheximide (**77**) was obtained as a white powder. LC-MS analysis revealed a molecular ion peak a m/z 282 for [M+H]$^+$ corresponding with the molecular formula $C_{15}H_{21}NO_5$ and five degrees of unsaturation. Analysis of ^1H and ^{13}C NMR data for **77**, along with 2D NMR data, indicated the presence of a similar glutarimidyl moiety with **75** and **76**, including two carbonyl carbons (δ_C 173.7, C-2 and C-6), two methylene groups (δ_C 37.2, C-3 and C-5; δ_H 2.56, 2H, and 2.26, 2H, H$_2$-3 and H$_2$-5), and one methine (δ_C 27.4, C-4; δ_H 2.22, H-4). Three out of five degrees of unsaturation being accounted for inferred that **77** contained an additional ring besides the glutarimide one. The COSY data of **77** exhibited a partial structure dimethyl-cyclohexanone, which was attached to the glutarimide moiety through C-7 and C-8. The NMR data is shown in Table 5.16.

Among all the compounds isolated, only nong-kang 101G (**76**) showed weak anti-BCG activity with MIC 343 μM.

In order to find out whether all of the constituents in the active fractions have been isolated, the NMR fingerprint between the 3 compounds and 4 active fractions were compared (Figure 5.31). Almost all the signals in active fractions can be also found in pure compounds, however, a few aromatic protons only existed in fraction 5 and 6, indicating the existence of additional analogs.

Table 5.16 ^{1}H and ^{13}C (600 MHz) NMR data of **75 – 77** in DMSO-d_6

| Position	75 δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC	76 δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC	77 δ_C, mult	δ_H, mult (*J* in Hz)	COSY	HMBC
1		10.71 s				10.78 s				10.58 s		
2	173.1 C				172.4 C				173.7 C			
3	47.38 CH$_2$	2.58 dd (3.1, 16.7)	3', 4	2, 4, 5, 7	40.2 CH$_2$	3.45 dd (5.4, 17.5)	3', 4	2, 4, 5, 7	37.2 CH$_2$	2.56 dd (4.5, 16.3)	3', 4	2/6, 4, 5, 7
		2.37 dd (10.4, 16.7)	3, 4	2, 4, 5, 7		3.02 dd (9.6, 17.5)	3, 4	2, 4, 5, 7		2.26 dd (6.2, 16.2)	3, 4	2/6, 4, 5, 7
4	26.4 CH	2.62 m	3/5, 3'/5', 7	2/6, 7	34.3 CH	2.59 m	3/5, 3'/5', 7	2/6, 7	27.4 CH	2.22 m	3/5, 3'/5', 7	2/6, 17
5	47.38 CH$_2$	2.58 dd (3.1, 16.7)	4, 5'	2, 3, 4, 7	40.2 CH$_2$	3.45 dd (5.4, 17.5)	4, 5'	2, 3, 4, 7	37.2 CH$_2$	2.56 dd (4.5, 16.3)	4, 5'	2, 3, 4, 7
		2.37 dd (10.4, 16.7)	4, 5	2, 3, 4, 7		3.02 dd (9.6, 17.5)	4, 5	2, 3, 4, 7		2.26 dd (6.2, 16.2)	4, 5	2, 3, 4, 7
6	173,1 C				172.4 C				173.7 C			
7	42.8 CH	3.16 d (6.6)	4	3/5, 4, 8	70.7 CH	4.03 dd (5.2, 10.4)	4,7-OH	3/5, 4, 8	40.5 CH	1.29 m	4, 8	3/5, 4, 8
8	205.7 C				197.0 C				65.3 CH	3.82 m	7, 9	4
9	118.7 C				121.6 C				50.9 CH	2.43 m	8, 14, 14'	8, 10
10	158.1 C				157.4 C				213.9 C			
11	126.4 C				125.9 C				40.2 CH	2.60 m	12, 12', 15	10, 12, 15
12	138.7 CH	7.28 s		10, 14, 15, 16	138.7 CH	7.28 s		10, 14, 15, 16	42.7 CH$_2$	1.78 m	11, 12', 13	10, 11, 13, 16
										1.47 m	11, 12, 13	10, 11, 13, 16
13	127.8 C				130.8 C				26.9 CH	2.06 m	12, 12', 14, 14', 16	
14	128.4 CH	7.58 s		8, 10, 12, 14	128.4 CH	7.58 s			35.4 CH$_2$	1.94 m	9, 13, 14'	8, 9, 10, 13, 16
								8, 10, 12, 14		1.61 m	9, 13, 14	8, 9, 10, 13, 16
15	15.5 CH$_3$	2.14 s		10, 11, 12	15.5 CH$_3$	2.14 s		10, 11, 12	14.9 CH$_3$	0.84 d (5.8)	11	10, 11, 12
16	20.3 CH$_3$	2.24 s		12, 13, 14	20.3 CH$_3$	2.24 s		12, 13, 14	18.6 CH$_3$	1,17 d (6.6)	13	12, 13, 14
7-OH						5.74 d (5.7)	7					
8-OH										4.34 m	8	
10-OH		12.20 s		9, 10, 11		12.20 s		9, 10, 11				

112
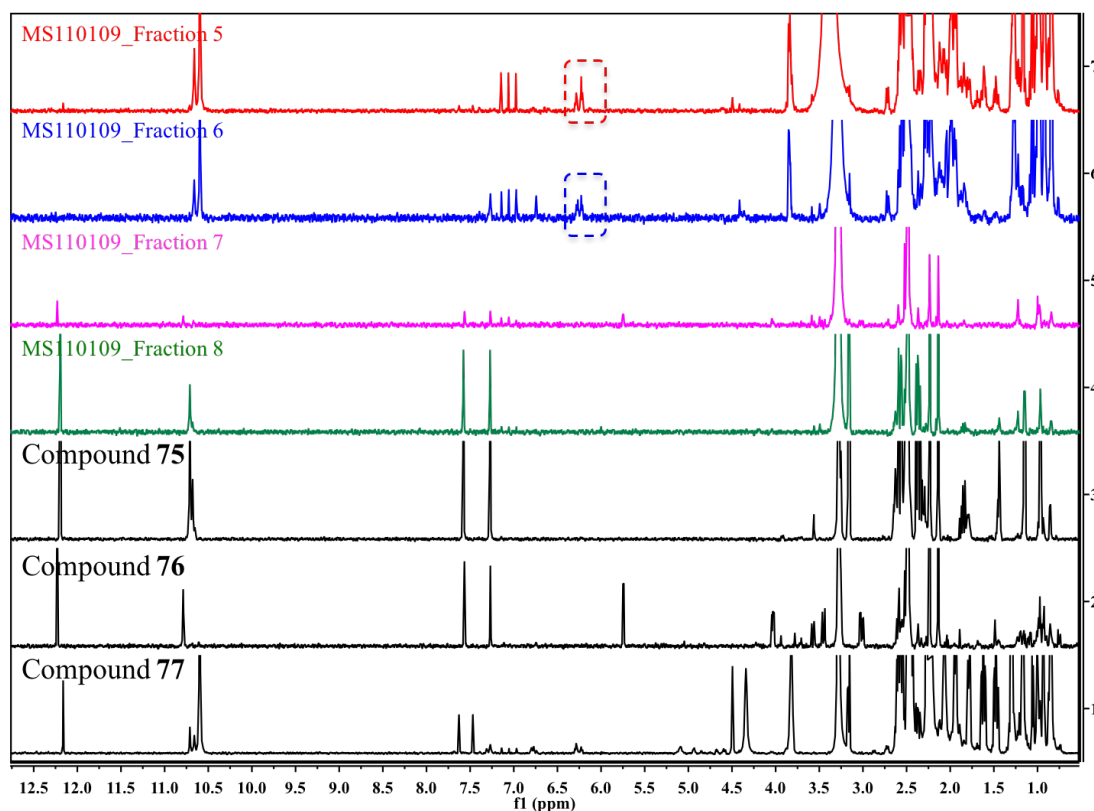
Figure 5.31 Comparison of NMR fingerprint between active fractions and compounds of MS110109_MPG.

5.4.6 Small-scale isolation of ES120055_AM2

The endophyte strain ES120055 was isolated from one Chinese Traditional Medicine *Saxifragaceae Astilbe* sample collected from Yunnan Province, China. The identification of ES120055 was performed based on the morphology and 16S ribosomal DNA (rDNA) phylogenetic analysis (Figure 5.32).
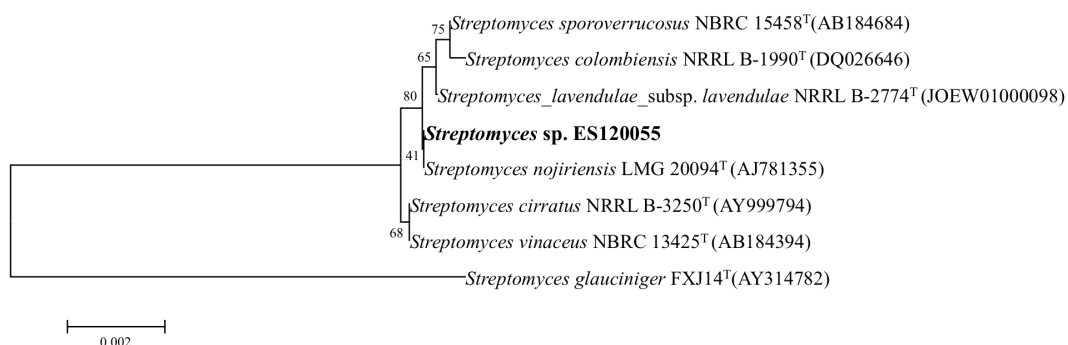
Figure 5.32 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of ES120055 and related species constructed by MEGA.
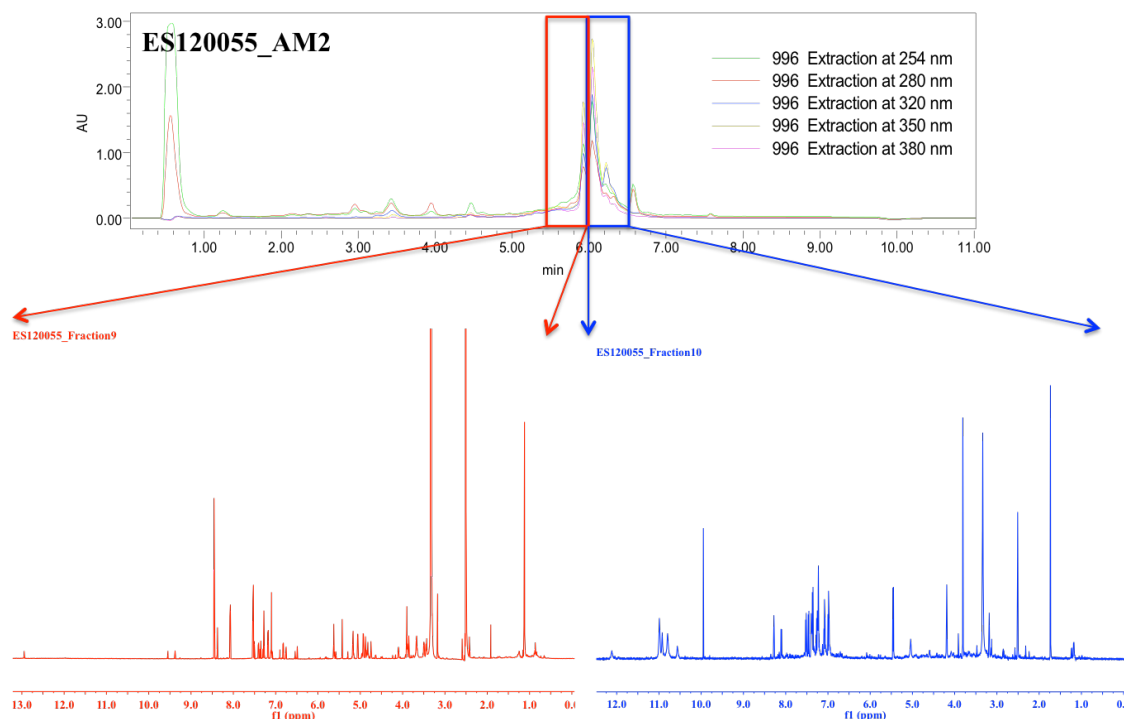
Figure 5.33 HPLC trace and NMR fingerprint of ES120055_AM2.

Cell-based screening on 11 fractions of ES120055_AM2 suggested only fraction 9 and 10 were active against BCG. ^{1}H NMR spectrum of 2 active fractions showed several coupled NMR signals in the aromatic region and exchangeable signals between 10-12 ppm (Figure 5.33).

Purification of active fraction 9 has identified two isoflavone natural products **78** and **79** (Figure 5.34), namely talosins B and daidzein 7-rhamnoside. The molecular weight of **78** was determined to be 562 on the basis of a LC-MS result and its molecular formula confirmed to be $C_{27}H_{30}O_{13}$ in combination with the ^{1}H and ^{13}C NMR data. The ^{1}H and ^{13}C NMR spectral data (Table 5.17) of **78**, together with the COSY, HSQC, and HMBC spectral data, revealed the presence of a 1,4- disubstituted benzene ring (δ_H 7.52; δ_C 130.2 and δ_H 7.10; δ_C 116.6), three isolated olefinic methines (δ_H 8.50; δ_C 154.7; δ_H 6.51; δ_C 99.9; δ_H 6.76; δ_C 95.1), two anomeric signals (δ_H 5.44; δ_C 98.6; δ_H 5.58; δ_C 98.7), eight oxygenated methanes (δ_H 3.85; δ_C 70.5; δ_H 3.43; δ_C 70.6 δ_H 3.66; δ_C 70.9; δ_H 3.49; δ_C 70.0), one singlet hydroxyl signal (δ_H 12.89), six doublet hydroxyl protons (δ_H 4.93, 4.81, 5.16, 4.75, 4.88, 5.06) one carbonyl carbon (δ_C 179.9), and seven sp2 fully subtituted carbons (δ_C 121.9, 105.8, 161.1, 161.0, 156.8, 123.8, 155.7). These data suggested that **78** was composed of an isoflavone and two sugar moieties. In the HMBC spectrum, the

olefinic proton at δ_H 8.50 (H-2) was coupled to the carbonyl carbon at δ_C 179.9 (C-4) and three sp2 fully substituted carbons at δ_C 156.8 (C-8a), 121.9 (C-3), and 123.8 (C-1′). The olefinic proton at δ_H 6.76 (H-8) was coupled to three sp2 fully substituted carbons at δ_C 156.8 (C-8a), 105.8 (C-4a), and 99.9 (C-6). Also, coupling was observed between the proton at δ_H 7.50 (H-2′) of the 1,4- disubstituted benzene ring and the carbon at C-3. The structure of the sugar was determined using COSY and HMBC experiments. In the HMBC spectrum, the anomeric proton at δ_H 5.44 (H-1″/H-1‴) was coupled to the oxygenated methine carbons at δ_C 70.6 (C-3″/C-3‴) and δ_C 70.0 (C-5″/C-5‴). In addition, couplings were observed from the methyl of H-6″ (H-6‴) to the carbons at C-4″ (C-4‴) and C-5″ (C-5‴). These spectral data showed the presence of two rhamnose-type structures. The linking position of the sugar moieties was determined using the HMBC data, which showed the cross peaks between H-1″(H-1‴) and C-7 and C-4′. Thus, the structure of **78** was determined as a known compound talosins B (Figure 5.34).

78: R_1= OH, R_2 = , R_3 =

79: R_1= H, R_2 = , R_3 = H

80: R_1= CH$_3$, R_2 = H, R_3 = CH$_3$
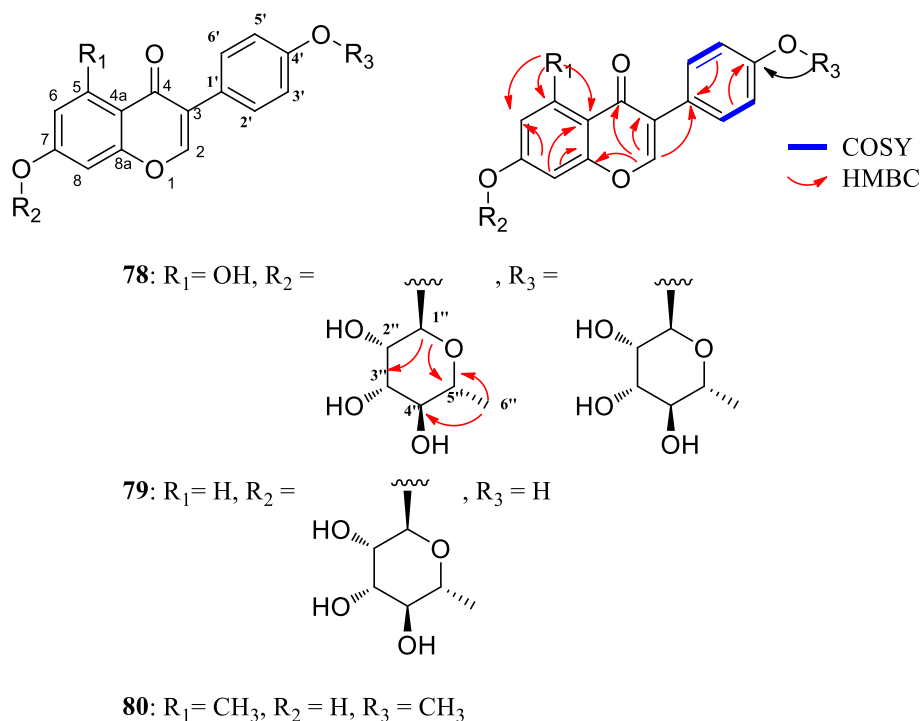
Figure 5.34 Structures and observed COSY and key HMBC correlations of **78**, **79** and **80**.

The molecular weight of **79** was determined to be 400 on the basis of a LC-MS result and its molecular formula as confirmed to be C$_{21}$H$_{20}$O$_8$ in combination with the ^{1}H and ^{13}C NMR data. The ^{1}H and ^{13}C NMR data of **79** were very similar to those of **78**. The major

Table 5.17 [1]H (600 MHz) and ^{13}C (600 MHz) NMR data of **78 – 80** in DMSO-d_6

Position	78		79		80		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
2	154.7 CH	8.50 s	154.3 CH	8.38 s	153.2 CH	8.40 s		3, 4, 8a, 1'
3	121.9 C		122.9 C		123.5 C			
4	179.9 C		180.5 C		175.3 C			
4a	105.8 C		106.3 C		114.3 C			
5	161.1 C		127.6 CH	8.07 d (8.95)	142.9 C			
6	99.9 CH	6.51 d (2.05)	115.7 CH	7.17 dd (2.41, 9.02)	114.5 CH	5.96 d (3.20)		4a, 5, 7, 8, 9
7	161.0 C		161.4 C		163.5 C			
8	95.1 CH	6.76 d (2.23)	104.0 CH	7.26 d (2.41)	98.14 CH	6.64 d (2.39)		4a, 6, 8a
8a	156.8 C		157.1 C		158.5 C			
9					22.9 CH$_3$	1.74 s		4a, 5, 6
1'	123.8 C		120.9 C		124.8 C			
2'/6'	130.6 CH	7.52 d (8.66)	131.1 CH	7.41 d (8.67)	130.1 CH	7.15 d (8.77)	3', 5'	3, 1', 2', 3', 4', 5', 6'
3'/5'	116.6 CH	7.10 d (8.91)	115.2 CH	6.82 d (8.65)	114.2 CH	6.81 d (7.38)	2' 6'	1', 2', 3', 4', 5', 6'
4'	155.7 C		157.6 C		159.8 C			
7'					55.8 CH$_3$	3.70 s		4'
1"	98.6 CH	5.44 d (1.54)	98.8 CH	5.62 d (1.87)			2"	7, 2", 3", 5"
2"	70.5 CH	3.85 m	69.8 CH	3.67 m			1", 3", 2"-OH	1", 3", 4"
3"	70.6 CH	3.43 m	65.0 CH	2.88 m			2", 4", 3"-OH	1", 5"
4"	70.9 CH	3.66 m	71.9 CH	3.44 m			3", 5", 4"-OH	2", 5", 6"
5"	70.0 CH	3.49 m	68.4 CH	4.10 m			4", 6"	1", 3", 6"
6"	18.4 CH	1.12 d (6.71)	16.4 CH	1.12 d (6.3)			5"	4", 5"
5-OH		12.89 s						4a, 6
2"-OH		4.93 d (5.72)		4.93 d (5.51)			2"	1", 3"
3"-OH		4.81 d (5.81)		4.82 d (6.15)			3"	2", 4"
4"-OH		5.16 d (4.32)		5.17 d (4.03)			4"	3", 5"
1'''	98.7 CH	5.58 d (1.54)					2'''	4', 3''', 5'''
2'''	70.5 CH	3.85 m					1''', 3''', 2'''-OH	4'''
3'''	70.6 CH	3.43 m					2''', 4''', 3'''-OH	1''', 4''', 5'''
4'''	70.9 CH	3.66 m					3''', 5''', 4'''-OH	2''', 6'''
5'''	70.0 CH	3.49 m					4''', 6'''	1''', 3''', 6'''
6'''	18.4 CH	1.12 d (6.44)					5'''	4''', 5'''
2"-OH		4.75 d (5.72)					3'''	1''', 3'''
3"-OH		4.88 d (6.27)					4'''	2''', 4'''
4"-OH		5.06 d (4.46)					5'''	3''', 5'''

differences between them in the ^{1}H and ^{13}C NMR spectra with regards to the HSQC data

were that one hydroxyl signal, one anomeric methine, four oxygenated methines, and one methyl signal were not observed in **79** (Table 5.17). These spectral data suggested that **79** is substituted by only one rhamnose at the C-7 hydroxyl group. In the COSY spectrum, the aromatic proton at δ_H 8.07 (H-5) was correlated with another aromatic proton at δ_H 7.17 (H-6), suggesting the hydroxyl group (5-OH) in **78** was replaced by a proton in **79**. Thus, compound **79** was determined as a known compound daidzein 7-rhamnoside (Figure 5.34).

The molecular weight of **80** was determined to be 282 on the basis of a LC-MS result and its molecular formula as confirmed to be $C_{17}H_{14}O_4$ in combination with the ^1H and ^{13}C NMR data. By comparison of NMR data with the above compounds, signals of sugar moieties were not found in **80**, additionally, one singlet methyl group (δ_H 1.74; δ_C 22.9) and one methoxy group (δ_H 3.70; δ_C 55.8) were observed (Table 5.17). Positions of the additional groups were confirmed by HMBC correlations from the methyl to carbons at δ_C 114.3 (C-4a), δ_C 142.9 (C- 5) and δ_C 114.5 (C-6), and from the methoxy group to carbon at δ_C 159.8 (C-4′). Structure of **80** is shown in Figure 5.34.

The absolute configuration of **78** and **79** were determined by comparing the $[\alpha]_D$ values in methanol with literature.[152, 153]

Chemical investigation of active fraction 10 was also conducted and one isoflavone compound and one small alkaloid have been isolated (Figure 5.35).
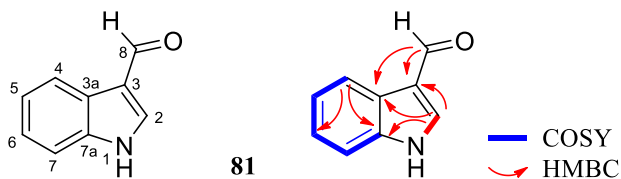
Figure 5.35 Structures and observed COSY and key HMBC correlations of **81**.

Compound **81** was obtained as a yellow solid. The LC-MS spectrum showed a molecular ion $[M+H]^+$ at m/z 146, suggesting the molecular formula C_9H_7NO. The ^1H NMR spectrum (Table 5.18) exhibited one series of aromatic signals and one exchangable proton. Analysis of the COSY spectrum indicated that signals at δ 8.08, 7.50, 7.24 and 7.20 formed a spin system that could be assigned to H-4, H-5, H-6 and H-7 of ring-A. Three singlets can be assigned to an indole proton (δ_H 8.26), one indole NH group (δ_H 12.09) and a methine group (δ_H 9.92) that appears in the ^{13}C NMR spectrum at unusually high field (δ_C 185.5), suggesting a formyl group. The structure was also confirmed by 2D

NMR (Table 5.18).

Table 5.18 ^{1}H (600 MHz) and ^{13}C (600 MHz) NMR data of **81** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1		12.09 s	2	
2	138.9 CH	8.26 s	1	3, 3a, 7a
3	118.6 C			
3a	124.5 C			
4	121.2 CH	8.08 dt (7.8, 1.0)	5	6, 7a
5	122.6 CH	7.20 td (7.4, 1.0)	4, 6	3a, 7
6	123.9 CH	7.24 td (7.6, 1.3)	5, 7	4, 7a
7	112.8 CH	7.50 dt (8.6, 1.0)	6	3a, 5
7a	137.5 C			
8	185.5 CH	9.92 s		3, 3a

Compounds **78** - **81** were found to be inactive against all the pathogens in our lab, including BCG, SA, MRSA, BS and SA.

Figure 5.36 showed the differences in NMR fingerprint of active fractions and identified compounds of ES120055. Almost every signal in fraction 9 was explained by compound **78** and **79**, however, it is clear that a couple of aromatic signals around 7 ppm and two doublets at 5.5 and 4.2 ppm in fraction 10 have not yet been isolated because of their low yield.
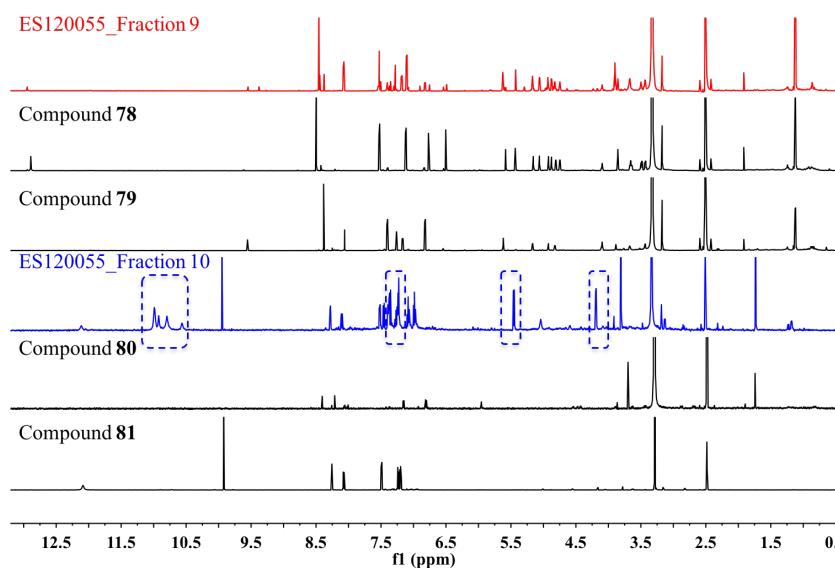
Figure 5.36 Comparison of NMR fingerprint between active fraction and compound of ES120055_AM2.

5.4.7 Small-scale isolation of ES120127_AM2

An endophyte ES120127 was isolated from TCM *Cirsium shansiense,* which was collected from Yunnan Province, China. It was identified as a *Streptomyces* sp. based on the morphology and 16S ribosomal DNA (rDNA) phylogenetic analysis (Figure 5.37).
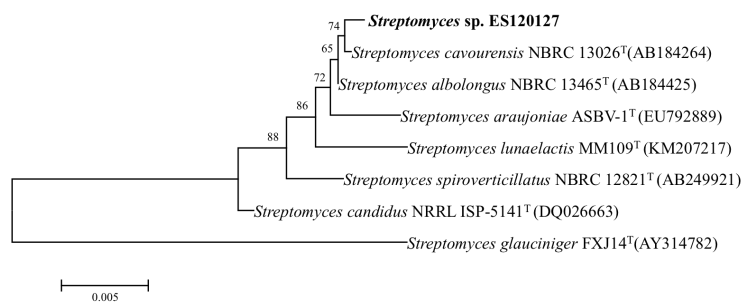
Figure 5.37 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of ES120127 and related species constructed by MEGA.

Fractions 7, 8 and 10 were shown to be active against BCG in the screening assay. From ^{1}H NMR fingerprint of the active fractions, several aromatic protons and methyl groups and exchangeable protons could be found in fraction 7, 8 and 10, respectively (Figure 5.38).
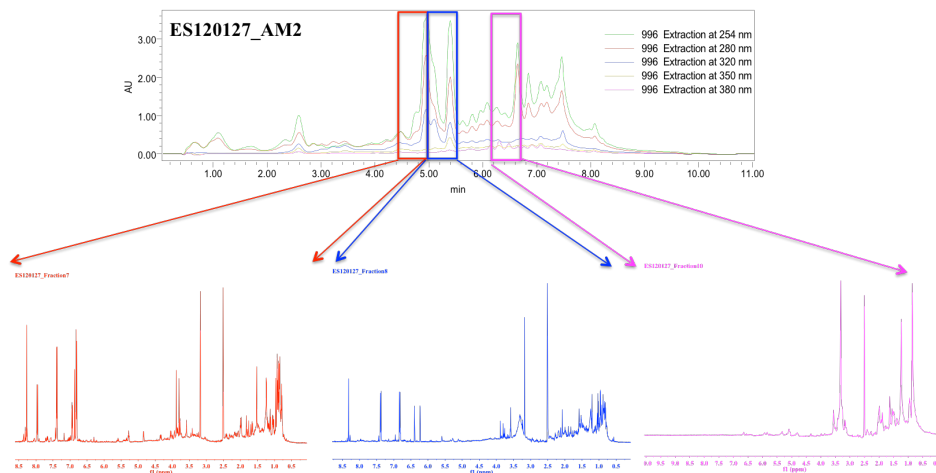
Figure 5.38 HPLC trace and NMR fingerprint of ES120127_AM2.

After small-scaled isolation, we identified two known isoflavones from fraction 7 (**82**) and 8 (**83**) (Figure 5.39) and a known depsipepide (**84**), which is L-O-Lac-L-Val-D-O-Hiv-D-Val from fraction 10 (Figure 5.40).

Compounds **82** and **83** were obtained as yellow powders and their molecular weights were determined as 254 and 270 by LC-MS and their molecular formulas established as

$C_{15}H_{10}O_4$ and $C_{15}H_{10}O_5$ with the analysis of ^1H and ^{13}C NMR data (Table 5.19). A comparison of the NMR data of **82**, **83** with those of **80** revealed that their NMR data were quite similar with the exception for the existence of a hydroxyl group in **82** and **83** instead of a methoxy group at C-4′ in **80**. Additionally, the methyl group at C-5 in compound **80** was replaced by a proton in **82** or a hydroxyl group in **83**, respectively. The additional hydroxyl in **83** was identified at C-5 via HMBC correlations from 5-OH (δ_H 12.96) to C-4a (δ_C 110.4), C-5 (δ_C 162.3) and C-6 (δ_C 99.5) (Figure 5.39).

82: R = H
83: R= OH
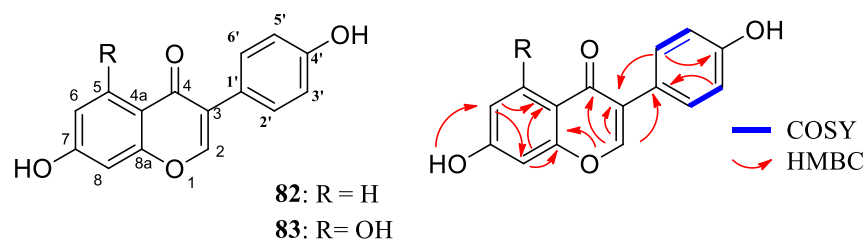
Figure 5.39 Structures and observed COSY and key HMBC correlations of **82** and **83**.

Table 5.19 ^1H (600 MHz) and ^{13}C (600 MHz) NMR data of **82** and **83** in DMSO-d_6

Position	82		83		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
2	153.2 CH	8.29 s	154.5 CH	8.32 s		3, 4, 8a, 1'
3	123.9 C		122.8 C			
4	175.2 C		180.5 C			
4a	117.1 C		110.4 C			
5	127.7 C	7.97 d (8.42)	162.3 C		6	7, 8a
6	115.6 CH	6.94 dd (1.96, 8.70)	99.5 CH	6.24 d (1.51)	5	4a, 7, 8
7	163.0 C		164.9 C			
8	102.6 CH	6.87 d (1.91)	94.2 CH	6.39 d (1.24)		4a, 6, 7, 8a
8a	157.9 C		158.1 C			
1'	123.0 C		122.6 C			
2'/6'	130.5 CH	7.39 d (7.70)	130.7 CH	7.39 d (8.34)	3', 5'	3, 2', 3', 4', 5', 6'
3'/5'	116.6 CH	6.81 d (8.66)	115.4 CH	6.83 d (8.62)	2' 6'	1', 2', 3', 4', 5', 6'
4'	157.6 C		157.9 C			
5-OH				12.96 s		4a, 5, 6
7-OH		9.51 s		9.58 s		6

Compound **84** was isolated as yellow solid and the molecular weights were determined as

388 by LC-MS. The ^{1}H NMR (600 MHz, in benzene-d_6, Table 5.20) spectrum of **84** exhibited 30 proton signals in total and there were three series of iso-propyl groups at δ_H 2.65 (m, H-3), δ_H 1.09 (d, J = 6.5 Hz, H-4), δ_H 0.86 (d, J = 6.6 Hz, H-5); δ_H 2.50 (m, H-8), δ_H 0.92 (d, J = 6.9 Hz, H-9), δ_H 1.02 (d, J = 6.8 Hz, H-10); δ_H 2.57 (m, H-13), δ_H 1.14 (d, J = 6.6 Hz, H-14), δ_H 0.85 (d, J = 6.6 Hz, H-15), one ethyl group at δ_H 1.57 (d, J = 6.8 Hz, H-18) and δ_H 4.06 (q, J = 6.8 Hz, H-17) and three downfield-shifted methines at δ_H 4.27 (dd, J = 10.7, 7.7 Hz, H-2), δ_H 4.42 (d, J = 3.3 Hz, H-7) and δ_H 4.13 (dd, J = 10.6, 6.4 Hz, H-12). The ^{13}C NMR spectrum of **84** (600 MHz, in Benzene-d_6, Table 5.3) showed 18 carbon signals, including four ester or amide carbonyl (δ_C 170.3, 171.0, 171.8 and 172.7), four sp$_3$ hydrolyzed carbons (δ_C 59.7, 78.7, 60.8 and 70.4), and seven methyel groups (δ_C 19.4, 19.7, 19.0, 16.7, 19.8, 19.4 and 17.3). The above NMR data were highly indicative of a peptide-like compound containing four amino acid residues. A combination of ^{1}H-^{13}C HSQC and ^{1}H-^{1}H COSY correlations (Figure 5.40 and Table 5.20) from H-2/H-3/H4, H-3/H-4, H-3/H-5, H-7/H-8/H-9, H-8/H-10, H-12/H-13/H-14, H-13/H-15, and H-17/H-18 suggested that **84** contained two valines (Val), one hydroxyl-isovaleric acid (Hiv) and one lactic acid (Lac). The presence of two valines was further supported by the correlations between the two NH signals and protons at δ_H 4.27 (dd, J = 10.7, 7.7 Hz, H-2) and δ_H 4.13 (dd, J = 10.6, 6.4 Hz, H-12). Sequencing the amino acid residues in **84** was accomplished by an HMBC experiment. HMBC correlations from H-18 to C-16 and C-17; from H-12 to C-16 and C-11; from H-7 to C-11 and C-6, and from H-2 to C-6 and C-1 established the connection of the four amino acid units. The absolute configuration of **84** was determined by comparing their specific rotation values in methanol with the literature.[154]
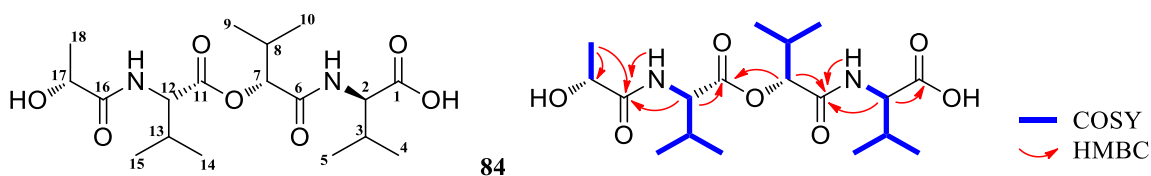
Figure 5.40 Structures and observed COSY and key HMBC correlations of **84**.

Compound **84** showed significant in vitro inhibitory activity against BCG with an MIC value of 8 μM. This is the first report of anti-TB activity for this known compound. Compounds **82** and **83** did not show any positive activities.

Table 5.20 [1]H (600 MHz) and ^{13}C (600 MHz) NMR data of **84** in benzene-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1	170.3, C			
2	59.7, CH	4.27 dd (10.7, 7.7)	3, 2-NH	1, 3, 6
3	28.2, CH	2.65 m	2, 4, 5	4, 5
4	29.4, CH$_3$	1.09 d (6.5)	3	3, 5
5	19.7, CH$_3$	0.86 d (6.6)	3	3, 4
6	171.0, C			
7	78.7, CH	5.42 d (3.3)	8	6, 10, 11
8	30.5, CH	2.50 m	7, 9, 10	
9	19.0, CH$_3$	0.92 d (6.9)	8	9, 10
10	16.7, CH$_3$	1.02 d (6.8)	8	8, 9
11	171.8, C			
12	60.8, CH	4.13 dd (10.6, 6.4)	13, 12-NH	11, 13, 14, 15, 16
13	28.3, CH	2.57 m	12, 14, 15	14, 15
14	19.8, CH$_3$	1.14 d (6.6)	13	13, 15
15	19.4, CH$_3$	0.85 d (6.6)	13	13, 14
16	172.7, C			
17	70.4, CH	4.06 q (6.8)	18	16, 18
18	17.3, CH$_3$	1.57 d (6.8)	17	16, 17
NH-12		8.34 d (6.6)	12	16
NH-2		8.41 d (7.9)	2	6
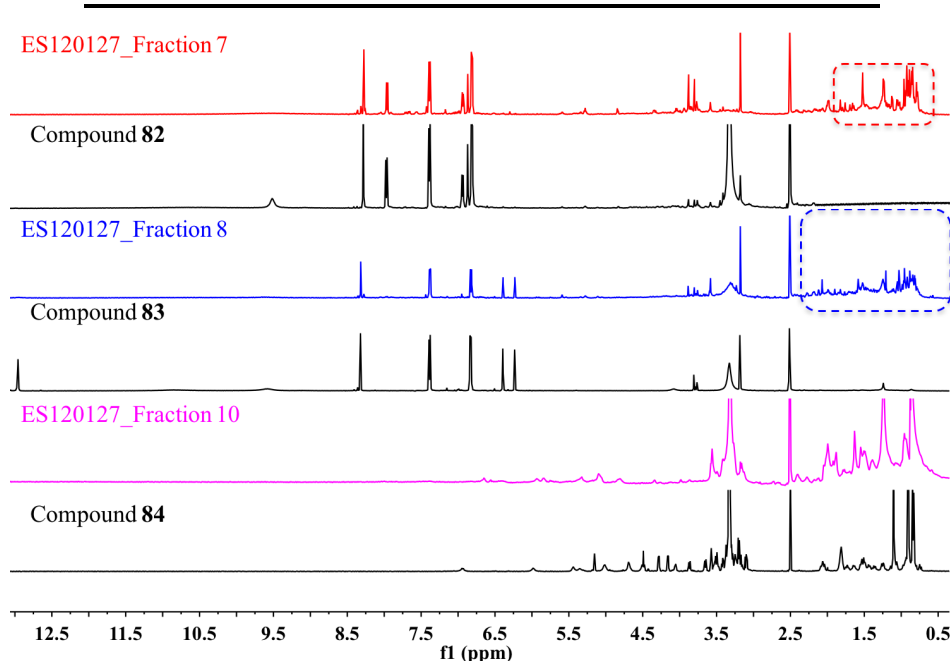
Figure 5.41 Comparison of NMR fingerprint between active fractions and compound of ES120127_AM2.

Based on the comparison between active fractions and isolated compounds, the majority

of protons in fractions belong to isolated compounds (Figure 5.41). A series of unidentified signals between 0.5-2.5 ppm in fraction 7 and 8 are probably from fatty acids, which represented the activity of fractions 7 and 8.

5.4.8 Small-scale isolation of other fractions

Apart from the above fractions, rapid chemical evaluation of active fractions derived from 20 crude extracts was also conducted. Figure 5.42 showed HPLC traces of two examples. Besides media constituents, no metabolite was identified from the studied fractions.
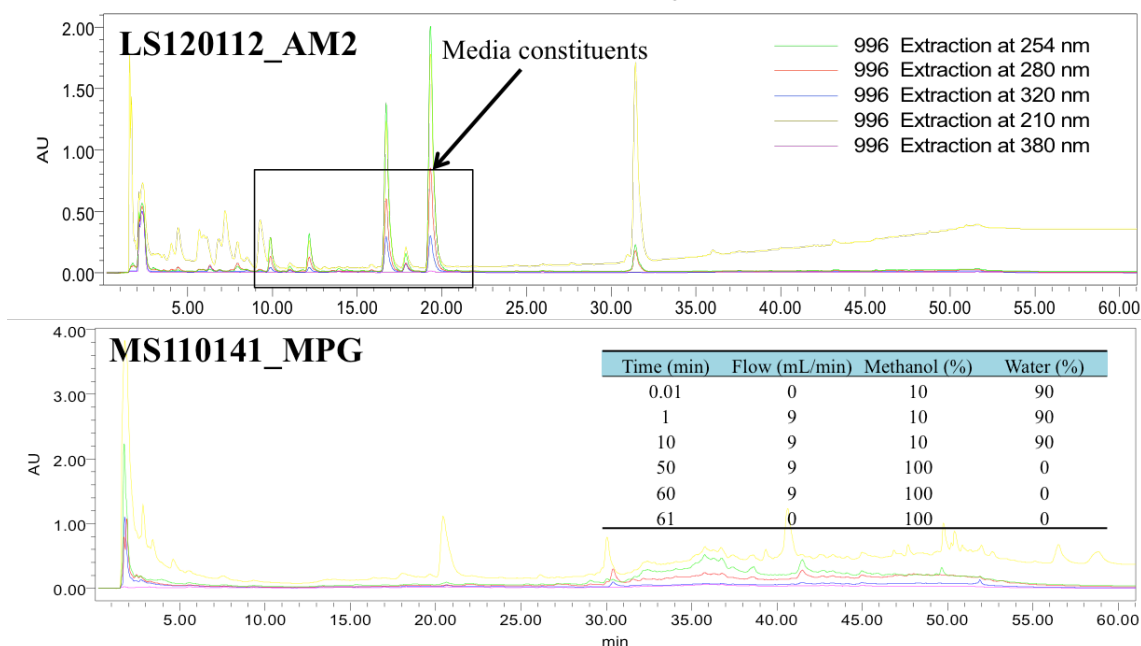
Figure 5.42 HPLC traces of 3 example active fractions. HPLC gradient was shown in the table.

5.5 Biological analysis of identified compounds

In this small-scaled investigation of active compounds from actinomycetes derived from unique environments project, 24 compounds sourced from 7 extracts were identified. All of the isolated compounds have been evaluated for their biological activities against *M. bovis* BCG, *S. aureus*, methicillin-resistant *S. aureus* (MRSA), *Bacillus subtilis*, and *Pseudomonas aeruginosa* (Table 5.21). Eleven natural products showed anti-BCG activities with MIC range from 1.5 to 343 μM. The most potent compound was compound **73**, which displayed promising activities against four Gram-positive bacteria with MIC values from less than 0.8 to 1.5 μM.

Table 5.21 Antimicrobial activities of compounds **61 – 84**

MIC (μM)	Organism (Strain)				
	BCG	SA	MRSA	BS	PA
61	6	6	6	<1.5	NA
62	25	50	50	25	NA
63	NA*	NA	NA	NA	NA
64	1.5	204	NA	51	NA
65	106	NA	NA	NA	NA
66	NA	NA	NA	NA	NA
67	14.15	28.3	28.3	NA	NA
68	27.9	27.9	27.9	NA	NA
69	30.7	30.7	NA	NA	NA
70	NA	NA	NA	NA	NA
71	NA	NA	NA	NA	NA
72	NA	NA	NA	NA	NA
73	1.5	<0.8	1.5	1.5	NA
74	49	98	98	49	NA
75	NA	NA	NA	NA	NA
76	343	NA	NA	NA	NA
77	NA	NA	NA	NA	NA
78	NA	NA	NA	NA	NA
79	NA	NA	NA	NA	NA
80	NA	NA	NA	NA	NA
81	NA	NA	NA	NA	NA
82	NA	NA	NA	NA	NA
83	NA	NA	NA	NA	NA
84	8	NA	NA	NA	NA
Control	0.37[a]	0.7[b]	0.7[b]	0.35[b]	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

BCG: Bacillus Calmette-Guérin (Pasteur 1173P2, BCG); SA: *Staphylococcus aureus* (ATCC 6538);

MRSA: Methicillin-resistant *S. aureus* (Clinical strain of Chaoyang hospital); BS: *Bacillus subtilis*

(ATCC 6633); PA: *Pseudomonas aeruginosa* (PAO1)

5.6 Conclusion

In conclusion, this work provided a successful strategy by a combination of HPLC

fractionation and ^{1}H NMR technique to rapidly recognize the known components and discover new natural products from microbial extracts. In total, 24 natural products were identified, including 11 active compounds and 1 new compound. Thus the developed strategy proved to be a very practical method to identify known and new metabolites from complex natural product extracts. In addition, it will also be useful in complex metabolomic analysis, especially in the fields of identification and characterization of the fingerprints and search for new components from complex natural resources.

However, as mentioned in the previous sections, several metabolites in the active fractions haven't been isolated due to their low yield. Or even no metabolite was identified from a number of active fractions because of the poor chemical diversity of those fractions or the low yielding of the natural products in those strains. Thus, methods to activate silent biosynthetic pathways are of major interest. The key issue for this approach should focus on the ways to induce or enhance the expression of cryptic or poorly expressed pathways to provide material for structure elucidation and biological testing.

Chapter VI NMR-OSMAC fingerprinting of 13 actinomycetes from unique environments

Abstract: The growing number of sequenced microbial genomes has revealed a remarkably large number of natural product biosynthetic clusters for which the products are still unknown. The aim of the present work was to apply a strategy to elicit new natural products by cultivating microorganisms in different fermentation conditions. The metabolomic analysis of 4160 fractions generated from 13 actinomycetes under 32 different culture conditions was carried out by ^{1}H NMR spectroscopy and multivariate analysis techniques. The principal component analysis (PCA) of the ^{1}H NMR spectra showed a clear discrimination between those samples by PC1 and PC2. The fractions with unique metabolites that are only produced under the specific conditions could also be outlined by PCA analysis. This method allows an efficient differentiation among a large dataset without any pre-purification steps. This work demonstrates the great potential of NMR spectroscopy in combination with metabolomic data analysis for the screening of large set of fractions. Based on the result, 37 outliers were identified as prioritized fractions and 6 of them were selected for further investigation.

6.1 Introduction

As a sequence of silent biosynthesis genes, the traditional methods to discover microbial natural products, which usually involve the collection and cultivation of strains, extraction, bioassay-guided isolation and structure elucidation, is often unsatisfactory. It is believed that many silent biosynthesis genes are only activated under specific conditions.[155] Several sequencing projects of different microorganisms have also confirmed that microbes have the potential to produce a lot more compounds from a single strain than previously known.[156-158] To gain access to this untapped reservoir of potentially bioactive structures, the biosynthesis of these putative metabolites needs to be induced. Even small changes in the culture medium may not only impact the quantity of a certain compound but also the general metabolic profile of an organism (Figure 6.1).[159]
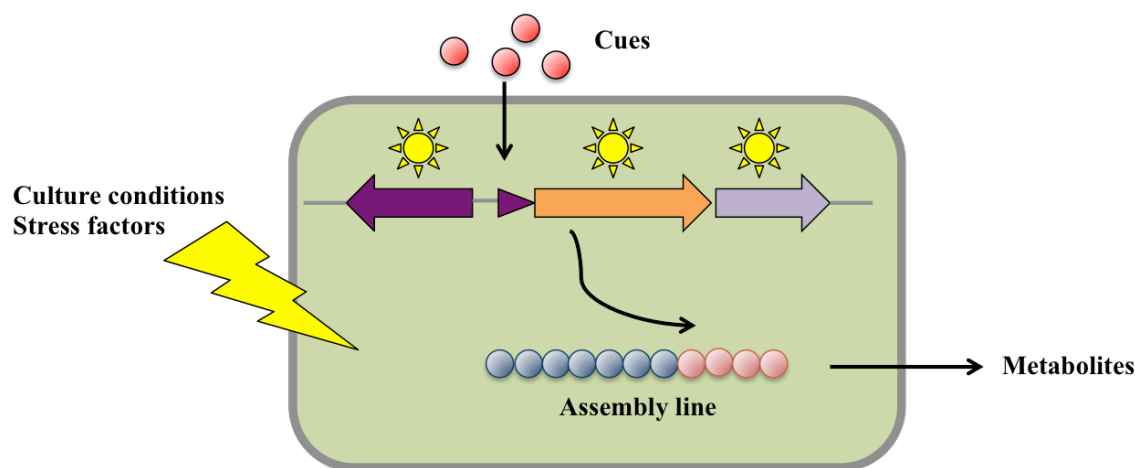
Figure 6.1 Influence of culture conditions, external (chemical) cues, and stress on natural product biosynthesis. (Modified from Scherlach K, et. al., 2009) [159]

In order to increase the number of secondary metabolites available from one microbial source, Zeeck and co-workers have investigated the systematic alteration of easily accessible cultivation parameters (for example, media composition, aeration, culture vessel, addition of enzyme inhibitors) on the production of secondary metabolites.[84] As a result, the OSMAC (One Strain - Many Compounds) approach was developed as a way of revealing nature's chemical diversity and to increase the number and yield of natural products from a single microbe.

6.2 Strategies

Historically, the increasing difficulty to find new chemical entities via high-throughput

screening has led to a decline in antibiotic research, while infectious diseases associated with multidrug resistance are spreading rapidly.[160] In this chapter, the chances of finding novel antimicrobials are increased, with focus on genetic, chemical, and biological methods to elicit the expression of biosynthetic gene clusters by the following aspects (Figure 6.2).

Endophytes   **Desert strains**   **Marine strains**

1. OSMAC strategies
2. Extraction
3. HPLC fractionation
4. NMR fingerprint
5. Data analysis

Unique metabolites
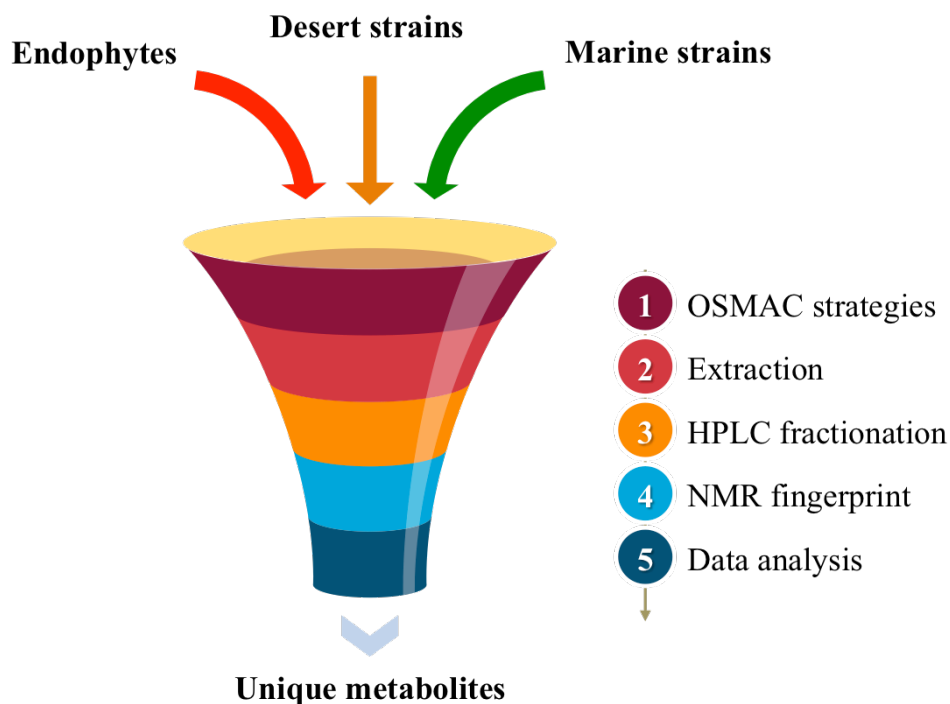
Figure 6.2. The flowchart of leads identification from unique actinomycetes based on OSMAC strategies.

First of all, with more and more chemically interesting and biologically significant metabolites isolated from microorganisms sourced from extreme environments due to the unique habitat, it has been shown that these microorganisms are rich sources of novel compounds with various biological activities.[161, 162]

Second, instead of using single cultivation condition, we applied OSMAC strategies with different cultivation parameters (for example, media composition, aeration, culture vessel, addition of enzyme inhibitors) to increase the number of secondary metabolites available from one microbial source.

Third, with the purpose to achieve lead-like enhanced extracts and to collect as much metabolites as we can, we explored readily available solid as well as liquid extraction methods.

Fourth, the method to generate lead-like enhanced (LLE) extracts and fractions allowed us to retain the lead- and drug-like constituents by selecting favorable physicochemical properties such as log P < 5.[163] As a result, fractions with low chemical complexity but high drug-like properties were collected.

Fifth, while the above technique examines the diversity of the drug-like natural product metabolome, it does not provide any information about the class of the small molecules in the fraction library. Herein, we detail a metabolic fingerprint procedure, based on the ^{1}H NMR spectroscopic analysis of the LLE fraction library, and demonstrate the effectiveness of the NMR-guided approach towards revealing the unique components of the drug-like natural product metabolome.

Last, unlike common dereplication methods that have to use every signal tested by chromatographic and spectroscopic methods to do database searching, we mainly focused on the outlier signals indicated by PCA results of OSMAC fractions. This is not only about saving time, but also directly targeting those unique compounds that are only produced under specific culture conditions.

6.2.1 Thirteen actinomycetes from unique environments
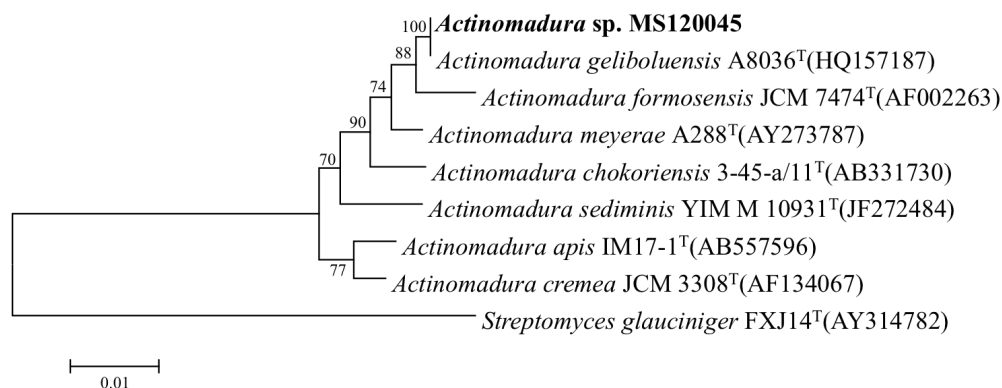
Figure 6.3 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS120045 and related species constructed by MEGA.

In this project, 13 actinomycetes from Chapter 5 were selected for further investigation, including 2 endophytes, 2 desert strains and 9 marine strains. Most of the strains belong to *Streptomyces* genus, only a marine strain MS120045 was identified as *Actinomadura* sp. according to 16s rRNA analysis (Figure 6.3). The 13 strains can also be divided into 2 groups according to the production of secondary metabolites under standard culture

condition (medium: AM2, culture period: 7 days). Five of them were found to produce at least one natural product under standard culture condition, however, besides medium compounds, no additional compounds or compounds with extremely low yield were found from the other eight marine strains (Table 6.1).

Table 6.1 Information of 13 selected actinomycetes

Strains	Produce NPs under standard condition
ES120055	Yes
ES120127	Yes
LS120167	Yes
LS120194	No
MS110104	No
MS110105	Yes
MS110109	Yes
MS110115	No
MS110142	No
MS110149	No
MS110158	No
MS110167	No
MS120045	No

6.2.2 OSMAC strategies

The OSMAC approach has resulted from the observation that very small changes in the cultivation conditions can completely shift the metabolic profile of many microorganisms.[164] Classically, systematic variations of culture media have been used to activate the production of additional secondary metabolites from microbes.[86] There is considerable evidence that some stress responses can also be used to trigger the expression of secondary metabolic genes: two that have been applied widely are heat shock and ethanol shock.[165] It was reported both stresses act through damage to the cell envelope and by driving the accumulation of misfolded or unfolded proteins.[166, 167] Moreover, varying additional cultivation parameters such as pH, oxygen supply, the addition of the enzyme inhibitors and inducers or the selection of drug-resistant bacteria mutants was also found to be effective in maximizing the diversity of microbial natural products.[86-88, 168, 169]

In this project, we applied 4 kinds of OSMAC conditions on 13 actinomycetes (Figure

6.4), which are variation of media constitutions (10 factors), different culture periods (2 factors), the use of stress conditions (10 factors) and the addition of inducer molecules (2 factors).

Figure 6.4 Four types of OSMAC strategies on 13 actinomycetes.

6.2.2.1 Culture media

In the present study, the metabolite profiles of the selected strains were investigated when cultured on different media. Media was selected under the consideration of their biological origin and produced secondary metabolites. To investigate if the strains squander the nutrients provided or require oligotrophic conditions to produce additional secondary metabolites, a rich-nutrient medium (medium R-1) and an oligotrophic medium (medium O-1) were selected. Several media used in the literature were also intentionally selected to induce the production of specific types of natural products, including terpene-induction media (media T-1, T-2, T-3, T-4 and T-5), alkaloid-induction media (media A-1 and A-2), polyketide-induction media (media PKS-1 and PKS-2), and borrelidin-induction medium (medium B-1). At last, a few widely used media for actinomycetes collected from different regions, including a medium for endophytes (medium E-1), 3 media for desert strains (media D-1, D-2 and D-3) and 5 media for marine strains (media M-1, M-2, M-3, M-4 and M-5) were selected. As a result, each strain was cultured in 10 different media (Table 6.2).

6.2.2.2 Culture periods

Two culture periods (7 days and 14 days) to each strain under different culture media

were conducted to investigate the effects of culture periods on the production of secondary metabolites.

Table 6.2 Ten different media for each strain

Strains	Medium 1	Medium 2	Medium 3	Medium 4	Medium 5	Medium 6	Medium 7	Medium 8	Medium 9	Medium 10
ES120055	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	E-1	D-1
ES120127	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	E-1	B-1
LS120167	R-1	O-1	T-1	T-2	T-3	T-4	T-5	B-1	D-2	D-3
LS120194	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	D-1	D-2
MS110104	R-1	O-1	T-1	A-1	PKS-1	M-1	M-2	M-3	M-4	M-5
MS110105	R-1	O-1	T-1	A-1	PKS-1	M-1	M-2	M-3	M-4	M-5
MS110109	R-1	O-1	T-1	A-1	PKS-1	M-1	M-2	M-3	M-4	M-5
MS110115	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2
MS110142	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2
MS110149	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2
MS110158	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2
MS110167	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2
MS120045	R-1	O-1	T-1	T-2	A-1	A-2	PKS-1	PKS-2	M-1	M-2

6.2.2.3 Stress conditions

Table 6.3 Information of 9 stress-inducing conditions and the normal condition

Strains	Temperature	pH	Ethanol
Stress-inducing condition 1	35 ℃	7.5	0
Stress-inducing condition 2	42 ℃	7.5	0
Stress-inducing condition 3	42 ℃ for 1 hour	7.5	0
Stress-inducing condition 4	28 ℃	3.5	0
Stress-inducing condition 5	28 ℃	5.5	0
Stress-inducing condition 6	28 ℃	9.5	0
Stress-inducing condition 7	28 ℃	7.5	1 mM
Stress-inducing condition 8	28 ℃	7.5	10 mM
Stress-inducing condition 9	28 ℃	7.5	100 mM
Normal condition	28 ℃	7.5	0

In addition to variation of culture conditions, the influence of various stress conditions on secondary metabolite biosynthesis has been enormous. The organisms were grown under normal and stress-inducing conditions for the generation of standard and perturbed metabolite profiles. Heat shock or raising the cultivation temperature has been shown as an effective way to produce bioactive natural products.[170, 171] It was reported that short-

time heat shock resulted in even higher natural products formation.[172] Another widely used parameter in stress-inducing conditions is the addition of different amount of ethanol into the culture medium. Microbial natural product biosynthesis can also be triggered by acidic pH shock.[173] Totally, 9 stress-inducing conditions were selected from the most well known stress conditions used to induce novel compounds (Table 6.3).

6.2.2.4 Inducing molecules

The expression of many secondary metabolite pathway genes can be controlled by regulators. For example, primary metabolites[174], enzyme inhibitors[175], pathway-specific regulators[176] and naturally produced chemical probes[177]. Quorum-sensing (QS) inducers were reported to have the potential to induce the production of natural products from *Streptomyces* sp..[87] Thus, in our project, we also selected two QS inducer molecules as chemical probes to regulate metabolites production by 13 actinomycetes (Figure 6.5).

L-Homoserine lactone hydrochloride   N-Carbobenzoxy-L-homoserine lactone
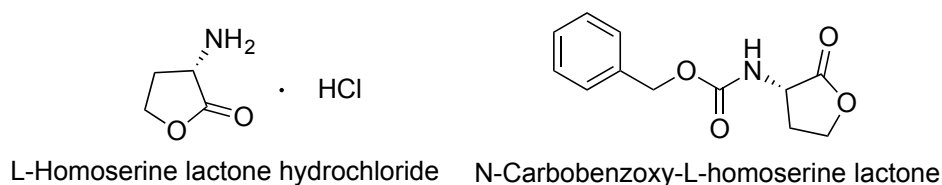
Figure 6.5 Structures of 2 selected QS inducer molecules.

6.2.3 Extraction
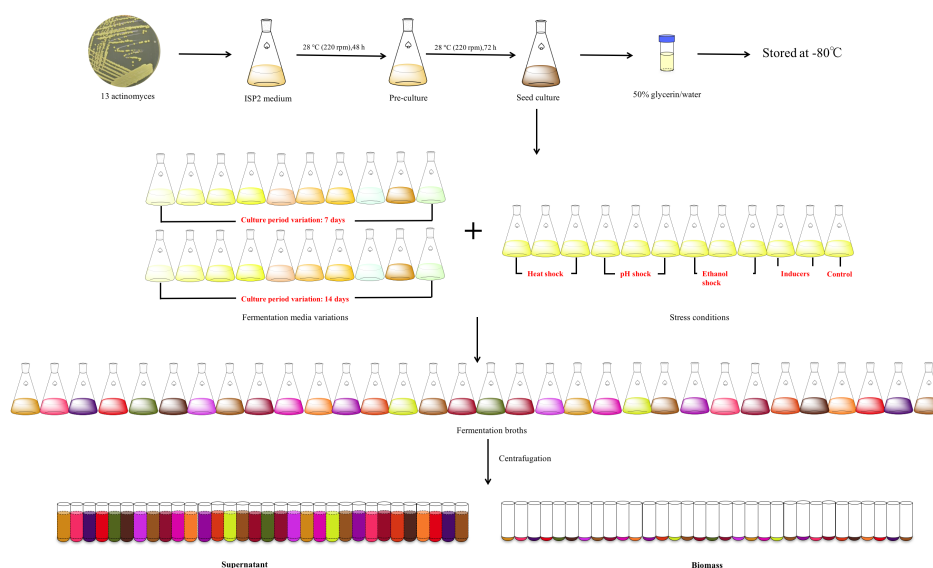
Figure 6.6 SOP for applying OSMAC strategies on 13 strains.

With the application of OSMAC strategies, each strain generated 32 broths (Figure 6.6). The two broths from duplicate flasks under that same conditions were combined and centrifuged to yield a supernatant and a mycelial cake of each condition. With the puropose of obtaining the most secondary metabolites produced by each strain under different conditions, multiple extraction methods were evaluated.

As reported previously, solid-phase extraction (SPE) with adsorbent DVB-NVP (HLB) was developed as an effective method to obtain crude extracts that contain components with log P < 5.[144] However, only a very limited amount of metabolites were extracted from a supernatant sample with this protocol (Figure 6.7).
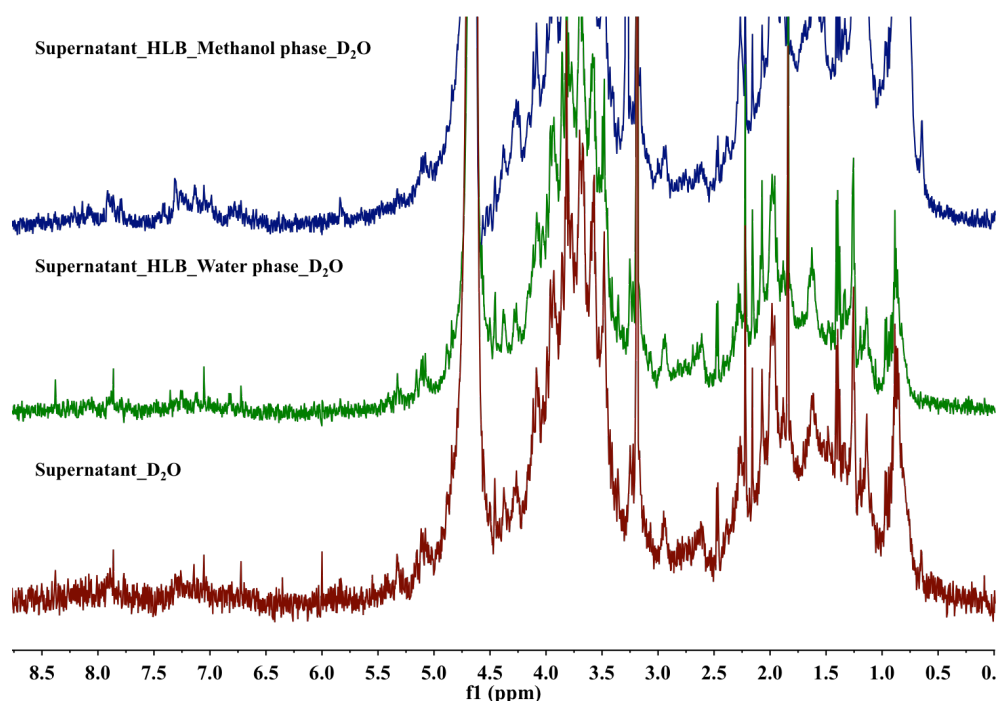
Figure 6.7 ^{1}H NMR spectra of supernatant crude extracts after HLB extraction.

One possible reason for the low recovery was considered to be the poor solubility of the supernatant samples since they had been directly dried from fermentation broths. To improve the solubility of samples, a liquid-phase extraction process before conducting HLB extraction was used. Supernatant sample was dissolved in 7.5 mL water and extracted three times with 7.5 mL n-butanol. A further HLB extraction was then applied and NMR spectra of crude extracts eluted from each step were shown in Figure 6.8. Nearly no differences were observed between two extraction processes, therefore, extraction of supernatant samples with n-butanol only was selected as the extraction

method in this project.
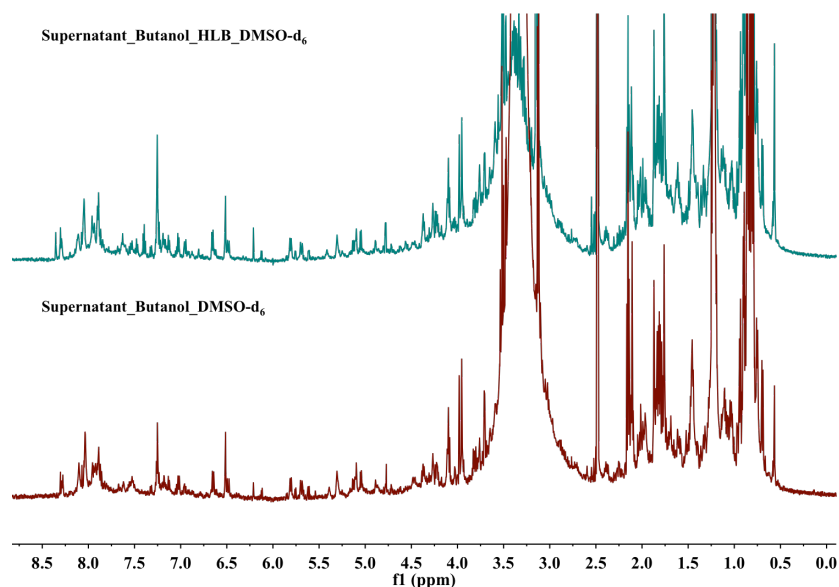
Figure 6.8 ^{1}H NMR spectra of supernatant crude extracts after n-butanol and HLB extraction.

The extraction effects of solid and liquid extraction methods on biomass samples were evaluated. The biomass samples were extracted with acetone (3 × 50 mL) and an additional n-butanol or HLB extraction process was then conducted (Figure 6.9). By comparison of the proton spectra of crude extracts derived from different extraction processes, no additional signals were collected from extraction with n-butanol or HLB. Extraction solvent of biomass samples was thus determined to be acetone.
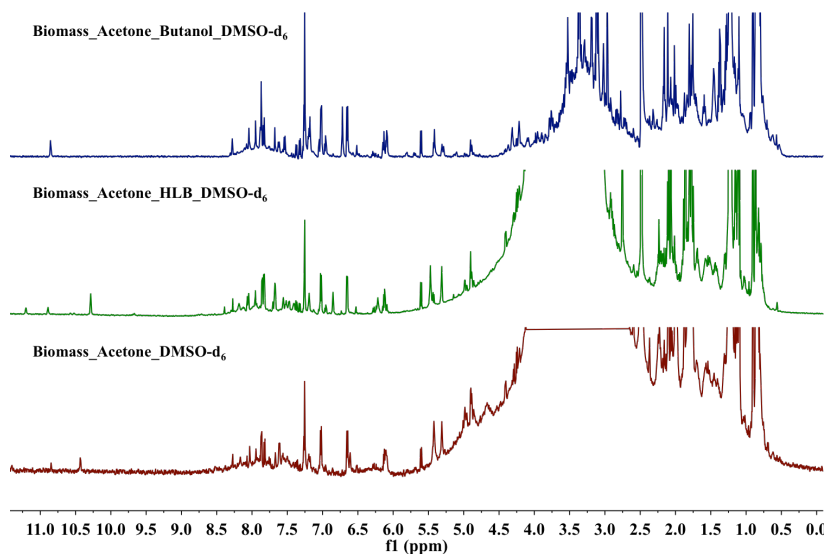
Figure 6.9 ^{1}H NMR spectra of biomass crude extracts after acetone, n-butanol and HLB extraction.

6.2.4 HPLC fractionation

LLE fractionation was accomplished according to the LLE protocol as followed. A small amount (10 mg) of dried samples were suspended in DMSO (300 μL). The extract (100 μL) was then fractionated by HPLC using a C_{18} Phenomenex Onyx monolithic HPLC analytical column (100 × 4.6 mm) with solvent conditions described in chapter 5. This fractionation provides a second log P filter allowing any remaining high log P components to be excluded. Also, to reduce the number of fractions and decrease the labour intensity in the following process, instead of 11 fractions, all the extracts were collected for 5 fractions every minute from 2 to 7 minutes. NMR spectra of 5 fractions from an example supernatant sample are shown in Figure 6.10. Comparing to crude extract, fractions contained only the mixtures of a small number of drug-like molecules.
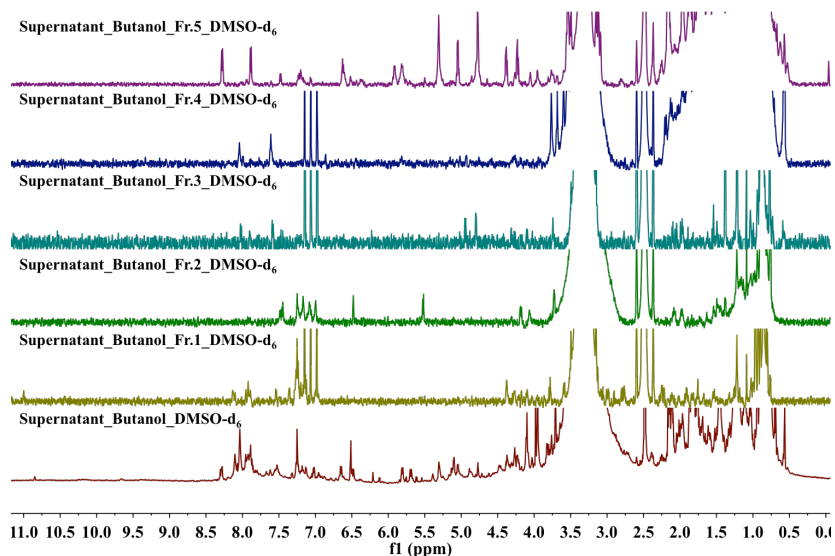
Figure 6.10 ^{1}H NMR spectra of supernatant crude extract and fractions.

6.2.5 NMR fingerprinting

Each LLE fraction was dissolved in 250 μL of deuterated DMSO and placed in a 5 mm NMR tube. The ^{1}H NMR experiments were performed on a Bruker (Rheinstetten, Germany) Ascend 800 MHz NMR. For each sample, the following parameters were applied, pw = 30°, p1 = 9.250 μs, d2 = 0 s, d1 = 1 s, at = 2.04 s, sw = 20.03 ppm, nt = 128 scans.. For quantitative analysis, peak integral was used.

6.2.6 Anti-TB cell-based assay

4160 LLE fractions derived from 13 actinomycetes under 32 OSMAC conditions by

HPLC fractionation of supernatant and biomass were evaluated for anti-BCG activity with a cell-based assay outlined in chapter 3.

6.2.7 Data analysis

The interesting regions of the NMR spectra for the analysis of the LLE fractions are from 0 to 15 ppm. The spectra were phase- and baseline- corrected manually by software Topspin. The ^{1}H NMR spectra were automatically reduced to ASCII files using AMIX (v.3.7, Bruker Biospin). Spectral intensities were scaled to the biggest peak (δ_H = 2.50 ppm) and reduced to integrated regions of equal width (0.02 ppm) corresponding to the region of 0.0 to 15.00. The region of 2.4 to 3.5 was excluded from the analysis because of residual signals of DMSO-$d5$ and HOD. The data set consisted of a 4160 × 696 matrix, in which rows represented the samples (4160 LLE fractions), and columns represented the 696 buckets of the ^{1}H NMR spectrum. Thus, each fraction was represented in the 696-dimensional space made of the 696 variables. Before analyzing the dataset, whether it can truly represent the NMR spectra needs to be confirmed. A digitalized spectrum with the data of all 696 buckets from one example fraction was constructed (Figure 6.11). Two spectra were perfectly matched, suggesting the high reliability of the bucket table generating process.
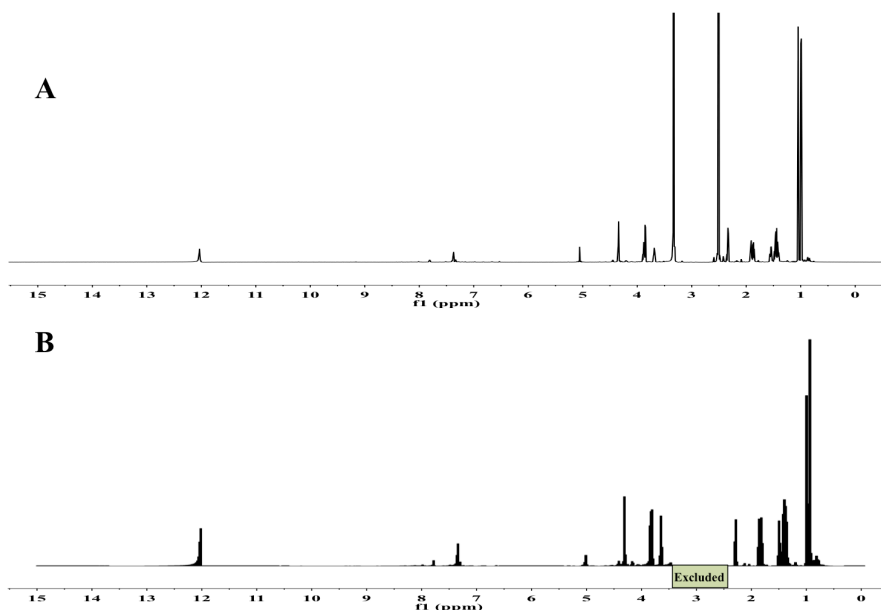
Figure 6.11 Comparison of 1H NMR spectrum with digitalized spectrum. A. ^{1}H NMR spectrum of an example fraction B. Digitalized spectrum of the example fraction with data in bucket table.

The whole spectrum was divided into 4 regions (0 - 2.4 ppm, 3.5 - 6 ppm, 6 - 10 ppm and 10 - 15 ppm) and analyzed separately to capture as many outliers as possible and decrease the disturbance by differences in intensities between saturated and unsaturated areas (Figure 6.12).

After a bucket table was calculated, PCA was applied by the same software AMIX. Because the variables (NMR chemical shifts) in this project are comparable, "No scaling" for the scaling of columns in the bucket table were selected. It preserves natural differences in intensities and it highlights dominant effects. To get rid of those columns in the bucket table that contain almost constant and small values would not produce a large variance and would not have a major influence on the PCA, the minimum variance level was set as 5%, which means variables that produce less than 5% variance of the maximum variance are removed from the PCA calculation. Confidence level was set as 95% to give score plot and loading plot.
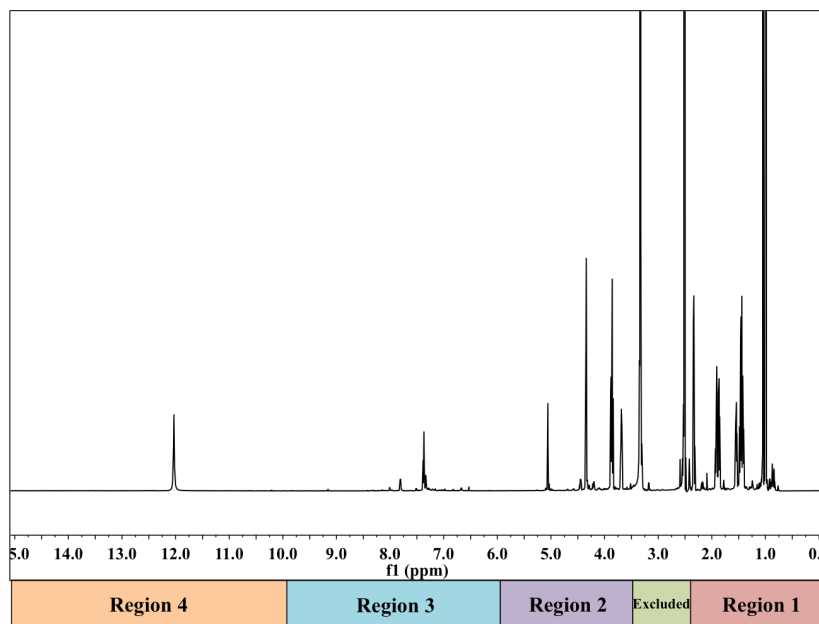
Figure 6.12 ^{1}H NMR spectra of an example fraction with different regions.

6.3 Results

6.3.1 Anti-TB HTS assay

Anti-TB activity of the library including 4160 fractions was evaluated in terms of BCG inhibition. In this study, all fractions were screened with an initial concentration of 50 μg/mL. Figure 6.13 represents the result of the preliminary screening test in heat map.
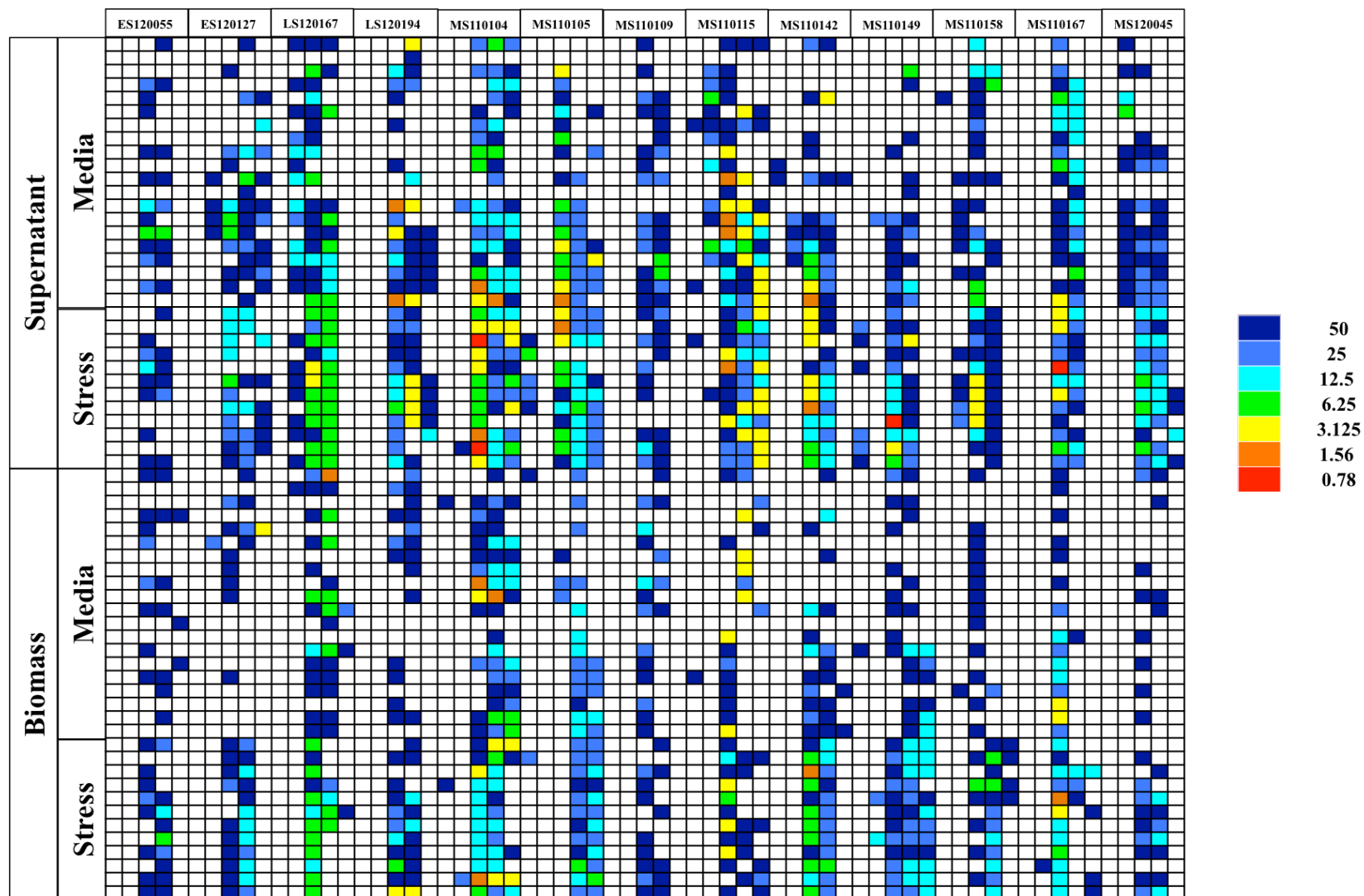
Figure 6.13 Heat map of the anti-BCG screening on 4160 LLE fractions. Each sample was presented in black blank.

The results of the anti-TB assay show that 1271 out of the 4160 (30.55%) fractions tested possessed biological activity ≤ 50 μg/mL. As compared to the positive and negative standards, 0.31% of the active fractions are very potent and inhibit the growth of BCG at 0.78 μg/mL (Figure 6.14). A further 1.49% and 6.29% of the active fractions were still active at concentractions of 1.56 and 3.125 μg/mL. In total, 8.09 % of the active samples can be regarded as potent BCG inhibitors with MIC values ≤ 3.125 μg/mL.
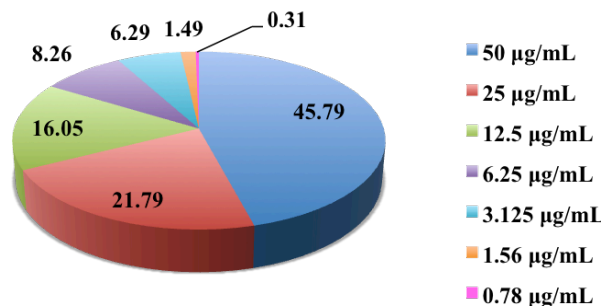
Figure 6.14 Pie chart of anti-BCG activities of 1271 active LLE fractions.

Within 2600 fractions generated from 520 crude extracts under different culture media and periods, 646 (24.84%) showed anti-BCG activity (Figure 6.15A). Comparing to fractions of biomass extracts, supernatant fractions exhibited a significantly higher hit rate. Regarding culture periods: longer culture period seems to produce more active fractions. Within 1560 fractions derived from 312 crude extracts of 13 strains under 12 stress conditions, 36.47% of fractions were active in anti-BCG screening, of which one third are contributed by pH shock fractions (Figure 6.15B).
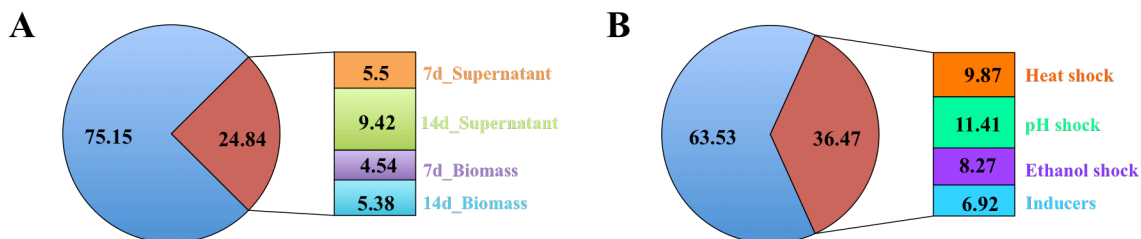
Figure 6.15 Pie chart of anti-BCG actives of 1271 active LLE fractions.

6.3.2 Data analysis

6.3.2.1 NMR intensity analysis

The effect of 32 OSMAC conditions on secondary metabolites production by 13 actinomycetes was studied. Total NMR intensity of each sample was calculated based on

its bucket table generated from the NMR fingerprint (Figure 6.16). In general, higher intensities were found in stress conditions compared to that of media variation conditions while longer cultivation time resulted in relatively higher NMR intensities, in other words, higher secondary metabolites yields. The highest NMR intensities were found in media 1, 4 and 6, which are all rich-nutrient media, in contrary; the oligotrophic medium (medium 2) gave the lowest NMR intensity among all conditions, suggesting essential effects of nutritional status on the selected actinomycetes.
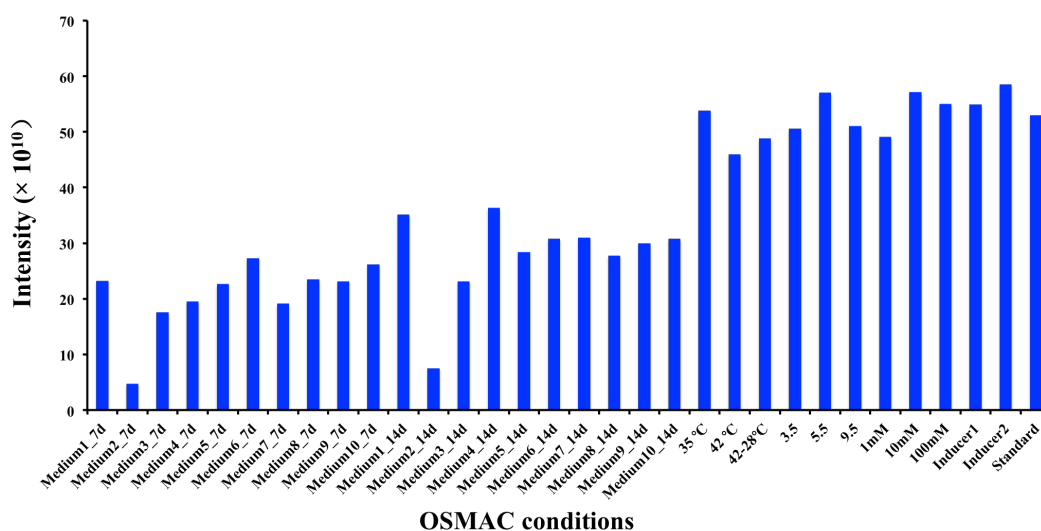
Figure 6.16 Distribution of total NMR intensities among 32 OSMAC conditions as an indicator of metabolite production.

Temperature is an important environmental factor for cell growth and product formation in the fermentation process.[178] When temperature increased from 28 °C (standard) to 35 °C, there was a slightly enhancement of secondary metabolites production (1.53% higher than standard). However, higher fermentation temperature (42 °C) caused the decrease effect of secondary metabolites production (13.23% lower than standard), which may be because of the inhibition of cell growth under high temperature. The decrease effect of 42 °C on secondary metabolites production was a little bit reversed by reducing the culture temperature back to 28 °C after 1 hour (7.84% lower than standard). It has been reported that high temperature stress could induce reactive oxygen species (ROS) in *Taxus yunnanensis*.[179] ROS induction could enhance the production of many metabolites.[180, 181] When pH decreasing from 7.5 (standard) to 5.5, great enhancement changes in metabolites production was observed (7.58% higher than standard). But more

acidic (3.5) or basic (9.5) pH shocks only resulted in lower productions, which could be due to the limited cell growth under those extreme conditions. Different addition concentrations of ethanol brought a change to the antibiotic production. Comparing to 1 mM and 100 mM ethanol, addition of 10 mM ethanol to strain cultures enhanced the metabolites production most (7.81% higher than standard). The effect of ethanol on metabolites biosynthesis may be due to its activity on changing the membrane structure, affecting steady-state growth and regulating related genes and carbon metabolism.[182] Both of two inducers lead to shifts on metabolites production, especially inducer 2 with a 10.47% higher NMR intensity than standard.

The effects of 32 OSMAC conditions on 13 selected actinomycetes were shown in Figure 6.17. The strain with the most significant changes in metabolites production is marine strain MS110149, with lowest NMR intensity in medium 2 (2.88×10^9) and high NMR intensities in medium 4 (9.28×10^{10}), 35°C (7.95×10^{10}), pH 5.5 (9.10×10^{10}) as well as inducer 2 (6.03×10^{10}). Comparing to media variation, stress conditions induced more changes to endophyte ES120055, marine strain MS110105 and MS110115, with peaks at ES120055_pH 5.5 (6.69×10^{10}), ES120055_inducer 2 (5.28×10^{10}), MS110105_42 °C (6.49×10^{10}), MS110105_pH 5.5 (7.13×10^{10}), MS110105_pH 9.5 (8.54×10^{10}), MS110105_EtOH 10 mM (8.48×10^{10}), MS110105_EtOH 100 mM (8.34×10^{10}), MS110105_inducer 1 (7.87×10^{10}), MS110105_inducer 2 (7.34×10^{10}), MS110115_35 °C (7.46×10^{10}), MS110115_pH 5.5 (6.90×10^{10}), MS110115_EtOH 1 mM (8.90×10^{10}) and MS110115_EtOH 10 mM (8.05×10^{10}), respectively. Land strain LS120194 showed great variations in NMR intensities under different media conditions, especially media 3 (3.45×10^{10}) and 4 (4.49×10^{10}). Strains showed enhanced NMR intensities by the effects of both media conditions and stress conditions, such as marine strain MS110104, with intensity values of 4.66×10^{10}, 5.30×10^{10} and 6.70×10^{10} for media 6, 10 and inducer 2, respectively.

The effects of 32 OSMAC conditions on the production of metabolites with structure characteristics in different NMR chemical shift regions were also examined (Figure 6.18). Consistent with expectations, NMR intensities decreased with the increasing of chemical shifts in each condition. The distribution patterns of each chemical shift region are almost the same with each other and are also very similar to those in Figure 6.16, where
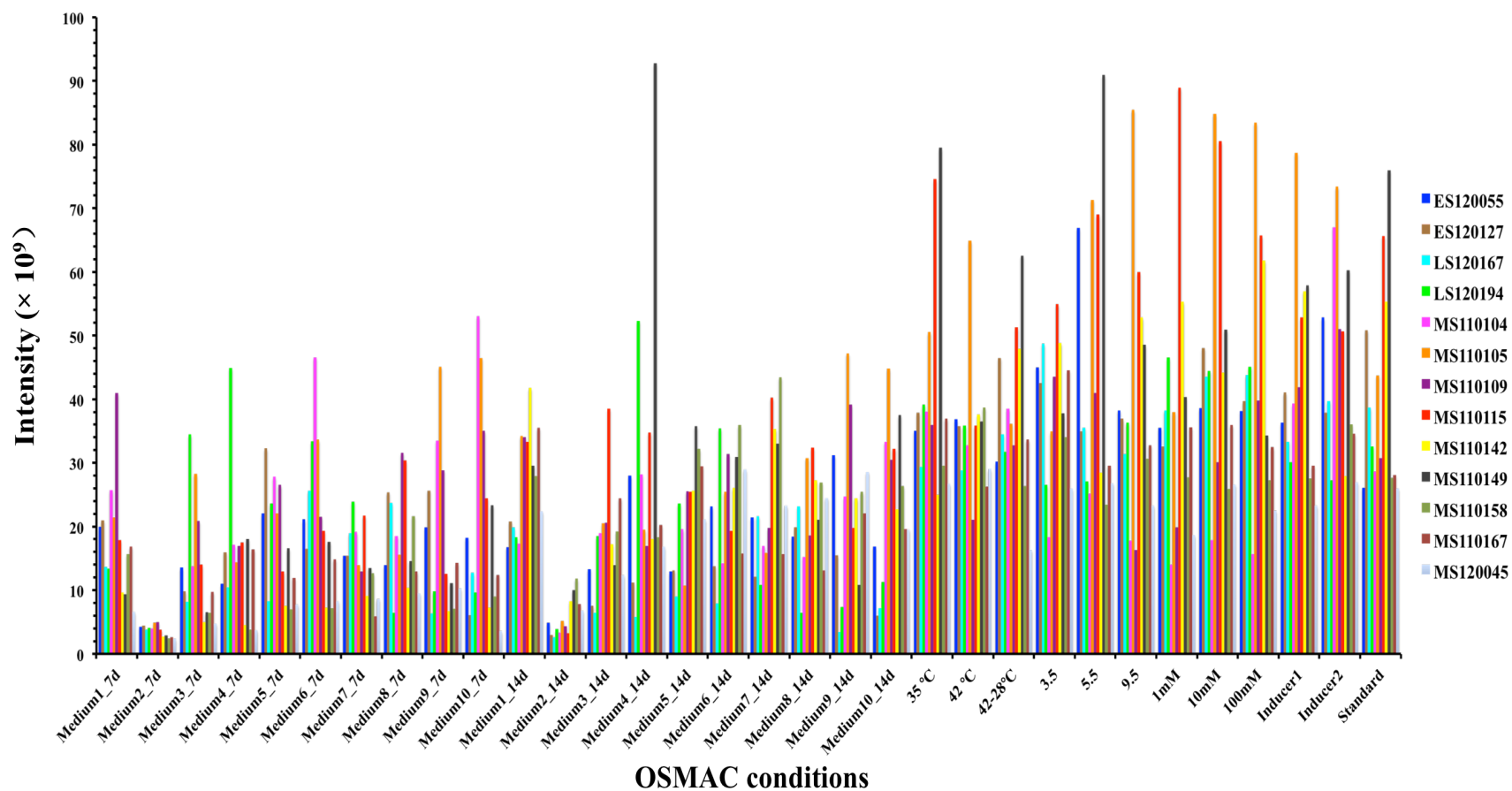
Figure 6.17 Distribution of total NMR intensities of 13 actinomycetes under 32 OSMAC conditions as an indicator of metabolite production.

chemical shifts in the whole spectra (0-15 ppm) were calculated. However, medium 6_14d (cultured in medium6 for 14 days) condition seems to produce more aromatic structures (6-10 ppm) comparing to its effect in producing structures with chemical shifts from 0-2.4 ppm and 3.5-6 ppm. Similar action was found in pH shock 3.5 (cultured at pH 3.5) condition, with NMR intensities 2.42×10^{11}, 2.07×10^{11}, 4.83×10^{10} and 8.45×10^{9} for different chemical regions 0-2.4 ppm, 3.5-6 ppm, 6-10 ppm and 10-15 ppm, respectively.
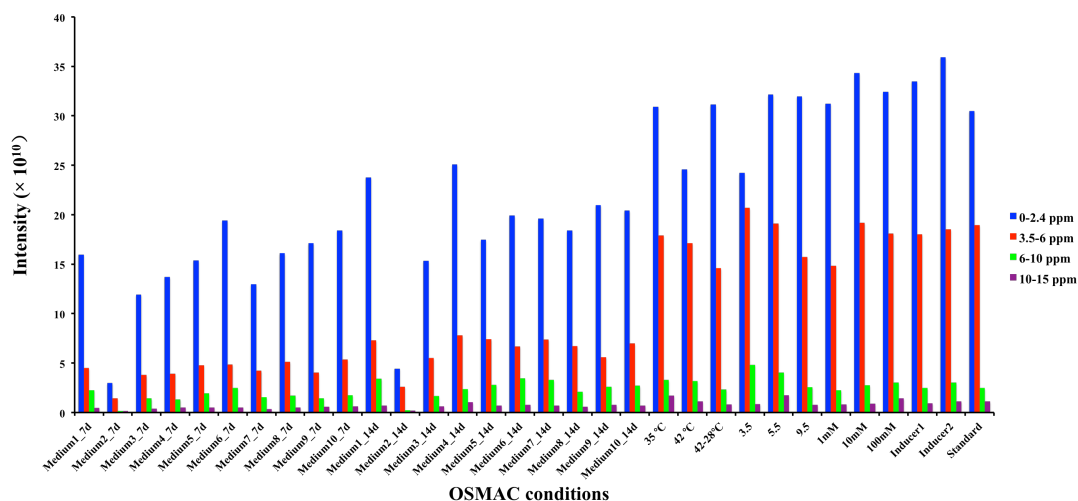
Figure 6.18 Distribution of total NMR intensities of different chemical shift regions as an indicator of metabolite production.
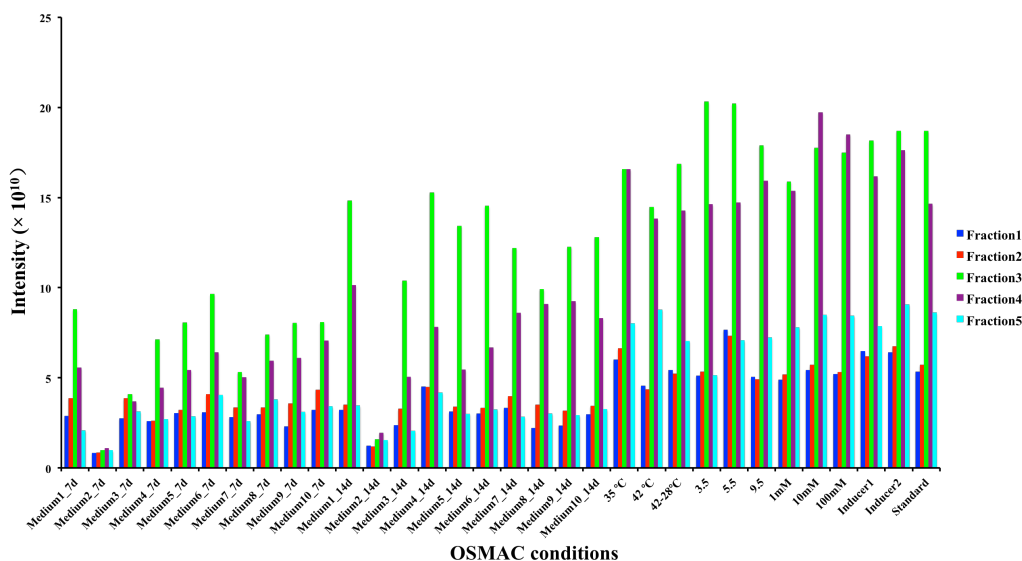
Figure 6.19 Distribution of total NMR intensities of different HPLC fractions as an indicator of metabolite production.

The effects of OSMAC conditions on biosynthesis of metabolites in different HPLC fractions by 13 actinomycetes were also investigated (Figure 6.19). In general, fraction 3 and 4 contained the most metabolites based on NMR intensities, and moreover, fraction 3 showed higher intensities than fraction 4 in most of the conditions, such as medium 3, 4, 5 and 6. However, higher intensities were found in fraction 4 rather than fraction 3 in condition 35 °C (cultured at 35 °C), 10 mM (cultured with the addition of 10 mM ethanol) and 100 mM (cultured with the addition of 100 mM ethanol).

6.3.2.2 Principal component analysis

PCA is an unsupervised method performed without using knowledge of sample class, and thus reduces the dimensionality of the data input while in a 2- or 3D map. By producing new linear combinations of the original variates, it plots data in order to indicate relationships between samples in a multidimensional space. It enables the easy comparison of microbe metabolic profiles. To analyze the data set obtained from 4160 fractions, the covariance method was applied. The main principal components (PC) to
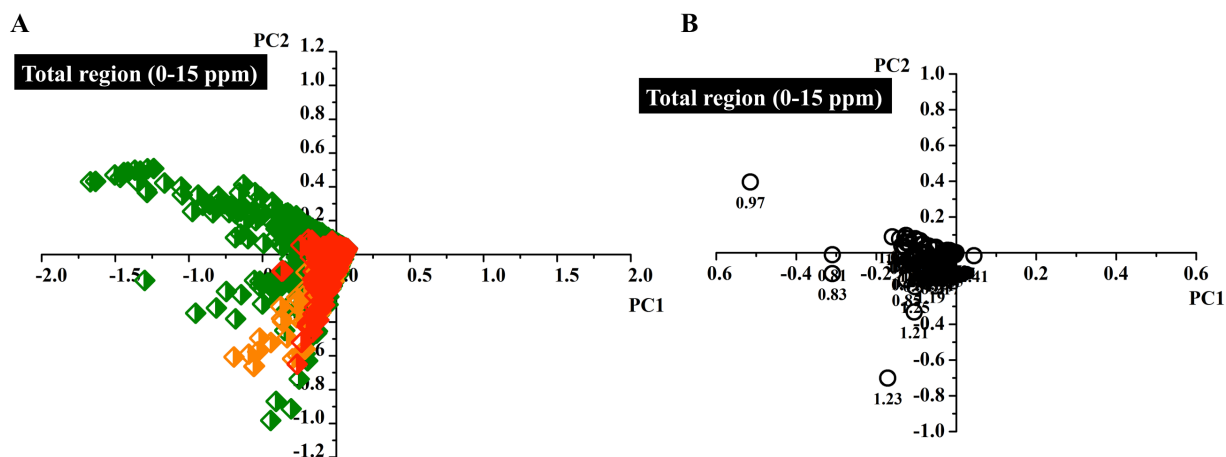
Figure 6.20 PCA results of 4160 fractions. A. Score plot of PC1 and PC2, fractions from endophytes, land strains and marine strains were shown in red, orange and green, respectively. B. Loading plot of PC1 and PC2, bucket values were shown as numbers below each circle.

differentiate the fractions are PC1 and PC2. In applying PCA to all the data, the fractions from different origins showed different cluster patterns (Figure 6.20A). Fractions from endophytes and land strains were clearly distinguished by PC1 and PC2, whereas fractions from marine strains were scattered over a much wider area. Therefore, it can be concluded that the metabolomic profiles of marine strains were considerably larger than

between endophytes and land strains under different OSMAC conditions. In the loading plot, it was found that PC1 and PC2 were mainly affected by high field signals such as from the signals at 1.23, 0.97, 0.81 and 0.83 (Figure 6.20B).

The buckets evaluated in this study were not completely analyzed because the PC values are dominantly affected by the strong signals in high field. The buckets with relatively low intensity in down fields are not sufficient to differentiate. For further metabolic analysis to obtain clear differentiation, the ^{1}H NMR spectra of the fractions were assessed for PCA in different chemical shifts regions (Figure 6.21).

No large differences were observed between the PCA results of the fractions in total region (0-15 ppm) (Figure 6.20 A and B) and region 1 (0-2.4 ppm) (Figure 6.21A and B), and almost all the spots in the score plot of region 4 (10-15 ppm) gathered in the same area (Figure 6.21G and H). Therefore the focus was placed on the results obtained from the other regions. As seen in Figure 6.21C and D (3.5-6 ppm), three fractions located in the third quadrant were found not to cluster with any of the others, suggesting the possibility of existence of unique metabolites. Visual inspection of the corresponding ^{1}H NMR spectra was carried out. However, only the fraction from the endophyte (ES120055_pH 5.5_Supernatant_fr.3) displayed various signals in saturated as well as aromatic regions (Figure 6.22). Not many interesting signals could be found in the other two outliers from marine strains, especially MS110149_Standard_Supernatant_fr.3. The NMR signals responsible for distinguishing them from the others are the huge peaks located around 4 ppm, which may be produced during the sample preparation process. This study has highlighted the importance of visual inspection of the ^{1}H NMR data after PCA, as not all outliers identified by PCA could be relied on to have unique metabolites signals.

The data set of aromatic region (6 – 10 ppm) examined by PCA analysis revealed five distinct outliers as shown in Figure 6.21 E and F. MS110158_Medium5_14d _Supernatant_fr.2 was well separated from the other outliers and it was highlighted because of the strong coupled signals at 7.27 and 7.31 ppm. Visual inspection of the other outliers in the second quadrant revealed different NMR signals at 7.79, 7.73, 7.71, 9.59 and 9.61 ppm (Figure 6.23).

Figure 6.21 Score plots and loading plots of PC1 and PC2 of 4160 fractions. Fractions from endophytes, land strains and marine strains in score plots were shown in red, orange and green, respectively. Bucket values in loading plots were shown as numbers below each circle. A. Score plot of region 1 (0 - 2.4 ppm). B. Loading plot of region 1 (0 - 2.4 ppm). C. Score plot of region 2 (3.5 - 6 ppm). D. Loading plot of region 1 (3.5 - 6 ppm). E. Score plot of region 3 (6 - 10 ppm). F. Loading plot of region 3 (6 - 10 ppm). G. Score plot of region 4 (10 - 15 ppm). H. Loading plot of region 1 (10 - 15 ppm).

Figure 6.22 ^{1}H NMR spectra of 3 outliers indicated by PCA results of region 2.

Only a few outliers were identified when analyzing 4160 fractions all together and only one of them (ES120055_pH 5.5_Supernatant_fr.3) showed promising characteristics in NMR spectrum. It was, hence, of interest whether more outliers with unique NMR signals could be resolved by PCA. This was investigated by re-subjecting the data sets generated from each strain separately to PCA analysis.

As with the previous evaluation, 13 data sets consist of 320 fractions from each strain were reanalyzed separately using the same software and parameters. The PCA generated score plots enabled distinction to be made among different OSMAC conditions in each strain. The metabolites responsible for distinguishing between the conditions are indicated in the loading plots. In score plots, 10 culture media conditions and 12 stress conditions were distinguished by different symbol shapes; samples from supernatant and biomass were shown in hollow and solid symbols, respectively; and fractions of different culture periods were differentiated with different fraction number forms: 1 - 5 (7 days) and 1′ − 5′ (14 days). Anti-BCG activity of each fraction was displayed in different color (Figure 6.24).

MS110105_Medium1_14d_Supernatant_fr.3

LS120194_Medium6_7d_Supernatant_fr.2

MS110142_Medium7_14d_Supernatant_fr.3

MS110167_Medium4_7d_Supernatant_fr.3

MS110158_Medium5_14d_Supernatant_fr.2
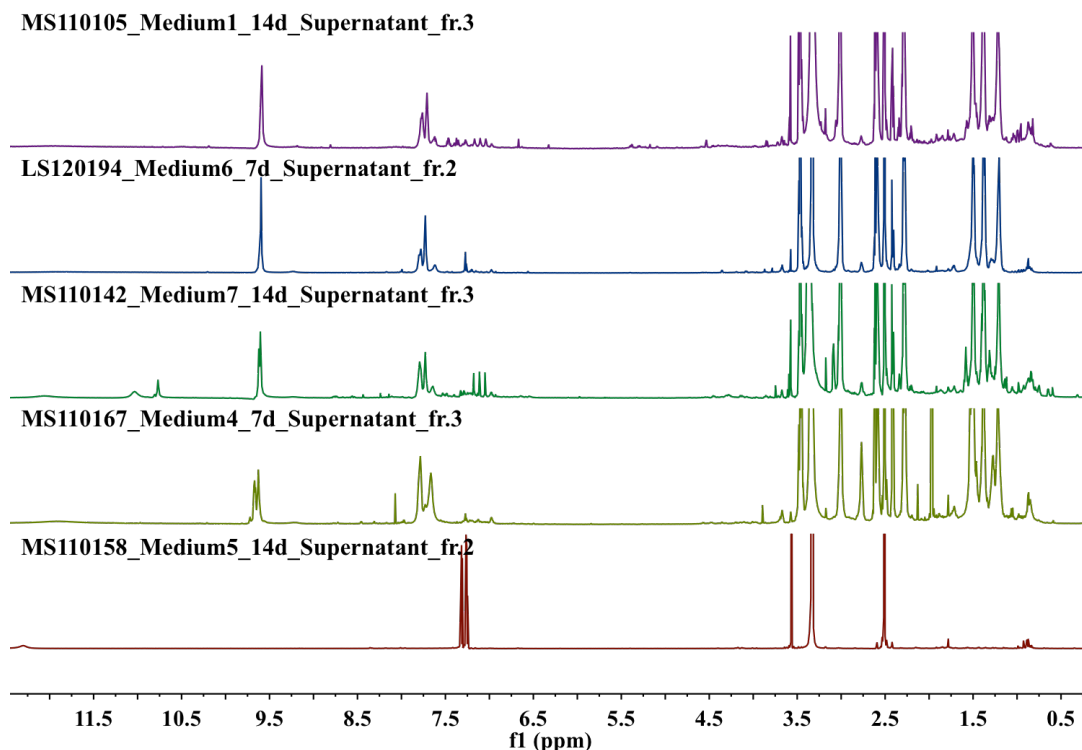
Figure 6.23 ¹H NMR spectra of 5 outliers indicated by PCA results of region 3.
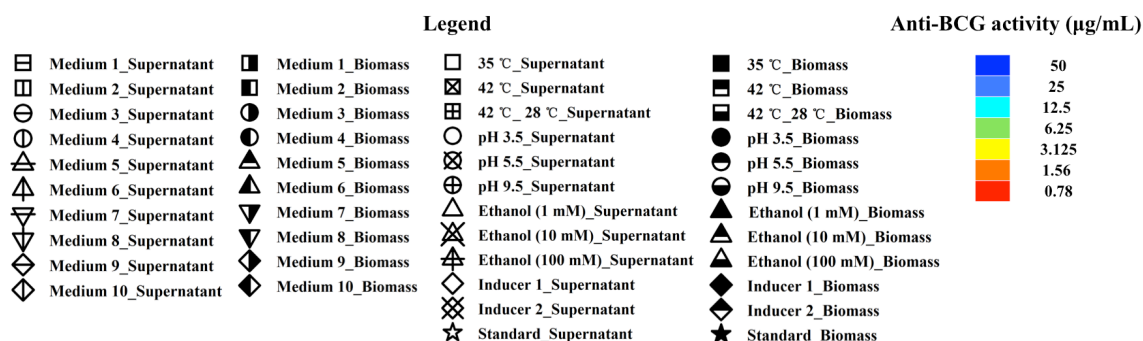
Figure 6.24 Legend information of score plot.

Outliers appeared in the first two principal components that cumulatively account for 79.30 % to 98.41 % of variation in each score plot were displayed in Table 6.4. In total, 156 outliers were identified from 52 score plots (13 actinomycetes × 4 regions) and 42 of them appeared more than once in different NMR spectral parts. The outliers were divided into 3 types according the visual inspection of their raw spectra. Type I outliers were highlighted mainly because of some huge signals generated in their spectra during the sample preparation or evaluation processes. As shown in Figure 6.25, the fraction 5 of MS110104_Medium6_7d_Biomass is totally separated from the other fractions by component 2 in the score plot. By examining the loading plot and the respective NMR

spectra (Figure 6.26), it is clear that the fraction is found to have a strong signal at 1.22 ppm but with no additional peaks comparing to the standard fraction.
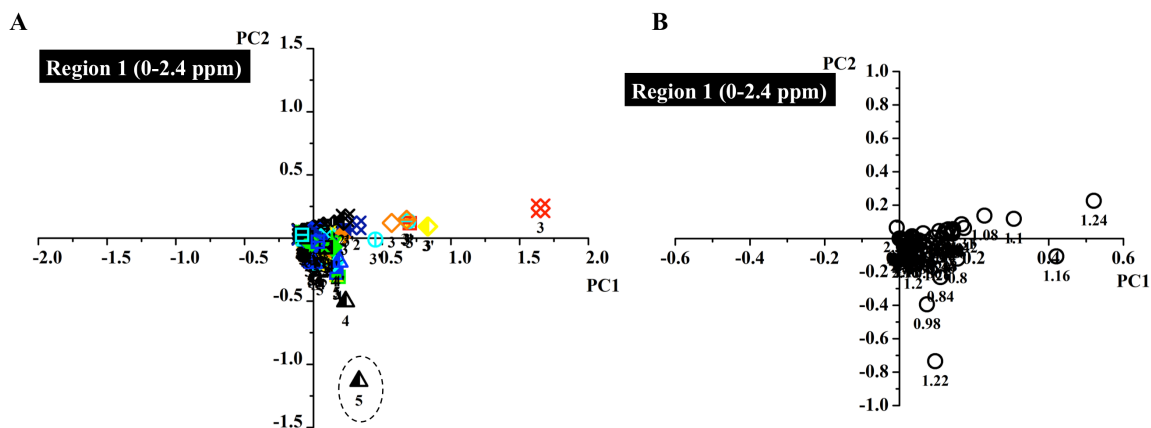
Figure 6.25 PCA results of 320 fractions from MS110104. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 - 2.4 ppm). Fraction numbers were shown below B. Loading plot of region 1 (0 - 2.4 ppm). Bucket values in loading plots were shown as numbers below each circle.
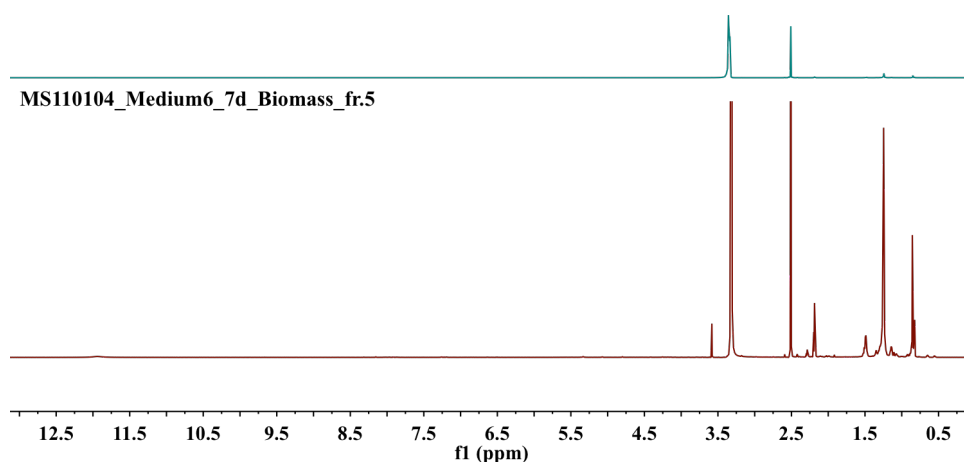
Figure 6.26 ^{1}H NMR spectra of the outlier of MS110104 indicated by PCA result of region 1. Bottom: MS110104_Medium6_7d_Biomass_fr.5, top: the corresponding standard fraction.

PCA were conducted based on bucket tables, which were intensities calculated from the selected NMR signals. Thus, PCA detects outliers according to the differences on

Table 6.4 Outliers identified from PCA results of 13 actinomycetes

Strains	Region 1 (0-2.4 ppm)	Region 2 (3.5-6 ppm)	Region 3 (6-10 ppm)	Region 4 (10-15 ppm)
ES120055	M9_14d_S_3, 5.5_S_3, M10_7d_S_2, 100 mM_B_4	5.5_S_3, Inducer2_S_4, 3.5_S_3	M9_14d_S_3, M1_7d_B_3	
ES120127	Standard_B_5, 100 mM_B_5	3.5_B_3, M1_14d_B_3, 42_B_5	M1_14d_B_3, M9_14d_S_3, 10mM_B_3, M9_7d_B_3, M8_14d_S_3	
LS120167	M4_14d_B_5, M1_14d_B_5, Standard_B_5, M1_14d_S_4	100mM_S_1	M6_7d_B_3, M8_14d_B_3, M7_14d_S_3, M1_14d_S_3, M8_14d_S_3, M8_7d_S_2	
LS120194	1mM_B_5, 10mM_B_5, 9.5_B_5, 9.5_B_4, 100mM_B_5, 1mM_B_4, 10mM_B_4, 10mM_S_3, 1mM_S_3	M4_14d_S_3, 10mM_S_3, 1mM_S_3, 100mM_S_3, 42_S_3	M4_14d_S_3, M6_7d_S_2, M3_7d_S_2, M6_14d_S_3, M3_14d_S_3	M4_14d_S_3
MS110104	Inducer2_S_3, M6_7d_B_5, M10_14d_S_3, M10_7d_B_3, 42_28_S_3, M3_14d_S_3, M9_14d_S_3, M9_7d_B_3, Inducer1_S_3, M4_14d_S_3	Inducer2_S_3, M2_14d_B_2, M2_14d_B_4, 42_S_3	42_B_4, 3.5_B_3, M6_14d_B_3	Inducer2_S_2, Inducer2_S_1, Inducer1_S_3
MS110105	M1_14d_S_3, M10_14d_S_3, M10_7d_S_2, M9_7d_B_4, Inducer1_B_4, M3_14d_S_4, M10_14d_S_4, Inducer2_B_4, M9_14d_B_4	Inducer1_B_4	M1_14d_S_3, M10_14d_S_3, M3_14d_S_4, M10_7d_S_2, M3_14d_S_3, M10_14d_S_4, M9_7d_B_3	
MS110109	M9_14d_S_3, M3_7d_S_5, M10_14d_B_4	Inducer2_S_3, 5.5_S_3, M9_14d_S_3	M9_14d_S_3, M1_14d_S_3, M5_14d_S_3, M4_14d_S_3, M9_7d_S_2, M1_7d_S_2	M9_14d_S_3, M9_7d_S_2, M1_7d_S_2
MS110115	M7_14d_S_3, M4_14d_B_4	M1_7d_B_4, M1_14d_B_4, M1_7d_B_5, M1_14d_B_5	5.5_B_3, 5.5_S_3, M1_14d_B_3, M1_14d_S_3, M6_14d_S_3, M7_14d_S_3	M7_14d_S_3, M10_14d_S_3, 5.5_S_3, 10mM_S_3, Inducer1_S_3
MS110142	M8_14d_S_3, 1mM_B_5	3.5_B_3, M8_14d_S_3, M6_14d_B_1	M7_14d_S_3, 42_S_3, 100mM_S_3, M8_14d_S_3	100mM_S_3, M7_14d_S_3

Table 6.4 (Continued)

Strains	Region 1 (0-2.4 ppm)	Region 2 (3.5-6 ppm)	Region 3 (6-10 ppm)	Region 4 (10-15 ppm)
MS110149	42_28_S_3, 5.5_S_3, M4_14d_S_3, 1mM_S_3, 9.5_S_3, Inducer1_S_3, 10mM_S_3, Inducer2_B_3, Inducer2_S_3, 100mM_S_3, M4_7d_S_3, 35_S_3, M4_14d_B_3	M4_14d_S_3	5.5_S_2, 35_S_2, 35_S_3, 42_S_5, 5.5_S_1, 35_S_1, M6_7d_S_3	5.5_S_2, 35_S_2, 35_S_3, 42_S_5, 5.5_S_1, 35_S_1, 9.5_S_3, Inducer2_B_3, Inducer2_S_3, Inducer1_S_3, 10mM_S_3, M4_14d_S_3, 1mM_S_3, 100mM_S_3, M4_7d_S_3
MS110158	42_B_5, M1_14d_B_5, M5_14d_B_5, Inducer2_B_5, M6_14d_B_5, 9.5_B_5	3.5_S_2, M5_14d_S_2	M5_14d_S_2, M6_14d_S_1, M6_14d_S_2, 42_S_3, 35_S_3, M7_14d_S_2, M9_14d_S_2, M10_14d_S_1	M5_14d_S_2
MS110167	M9_14d_B_5, M4_7d_S_3, M1_7d_S_3, 5.5 S_3	M9_14d_B_5, M1_7d_S_3, 5.5 S_3	M4_7d_S_3, M1_7d_S_3, 3.5_S_3	M9_7d_B_5, M4_7d_S_3
MS120045	M8_14d_B_5, Standard_B_5, 10mM_B_4, M9_14d_B_5, M9_14d_S_4, M8_14d_S_4	M1_14d_S_3, M6_14d_S_3, 1mM_B_4	M1_14d_S_3, M6_14d_S_3, M7_7d_S_3, M9_14d_S_3	M1_14d_S_3, M6_14d_S_3

Same fractions are shown in same color.

intensities, including samples containing the same metabolites with the others but in a higher quantity (type 2 outliers), or samples containing new generated metabolites from the others (type 3 outliers). An example of outlier that displayed the effect of increasing the metabolites production was shown in Figure 6.27. The metabolites production in endophyte ES120127 was enhanced by culturing in medium 9 for 14 days comparing to the standard condition and same medium for 7 days (Figure 6.28).
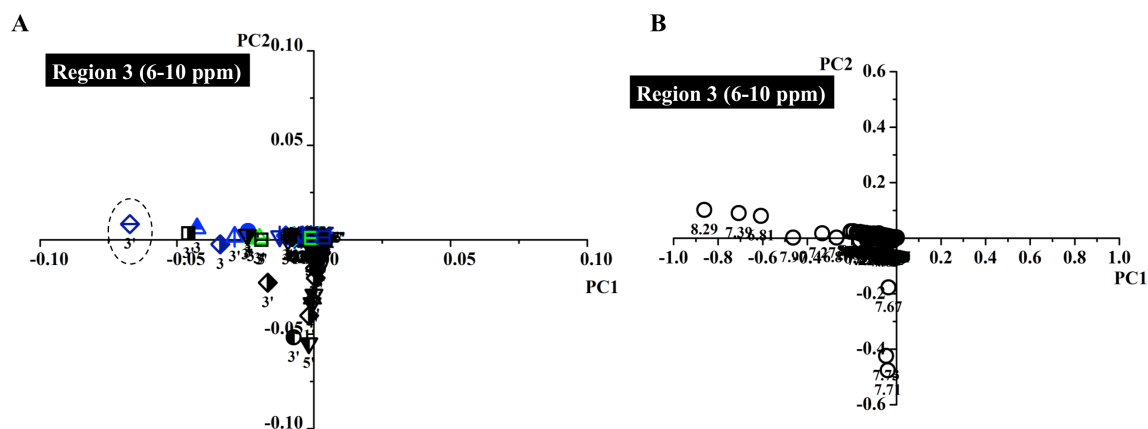
Figure 6.27 PCA results of 320 fractions from ES120127. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 3 (6 - 10 ppm). Fraction numbers were shown below each symbol. B. Loading plot of region 3 (6 - 10 ppm). Bucket values in loading plots were shown as numbers below each circle.
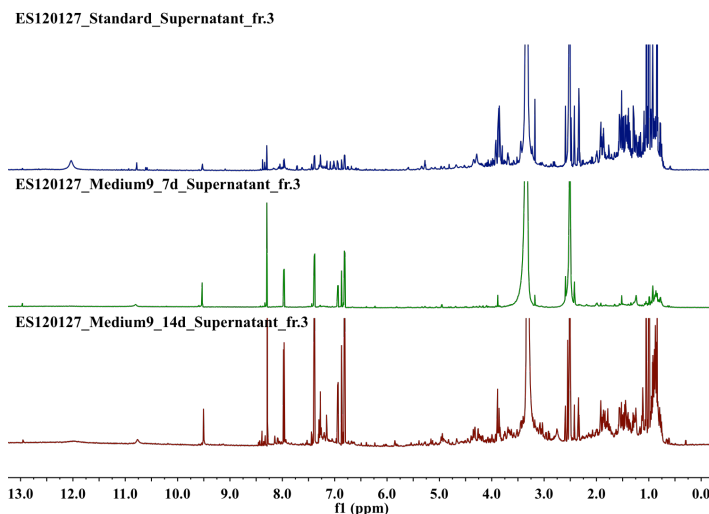
Figure 6.28 ^{1}H NMR spectra of the outlier of ES120127 indicated by PCA result of region 3. Bottom: ES120127_Medium9_14d_Supernatant_fr.3, middle: ES120127_Medium9_7d_ Supernatant_fr.3, top: the corresponding standard fraction.

Apart from culture media variation, the addition of ethanol in standard medium was
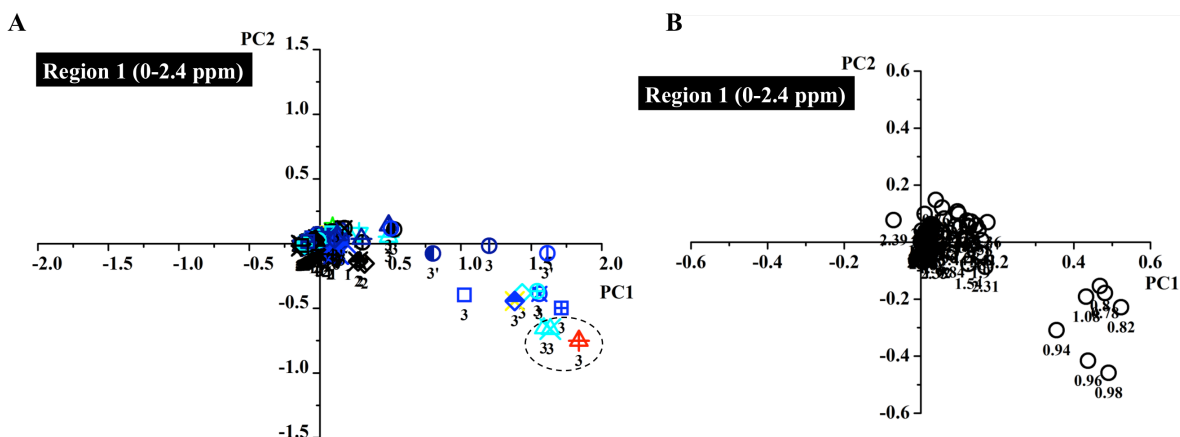
Figure 6.29 PCA results of 320 fractions from MS110149. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 – 2.4 ppm). Fraction numbers were shown below each symbol. B. Loading plot of region 1 (0 – 2.4 ppm). Bucket values in loading plots were shown as numbers below each circle.
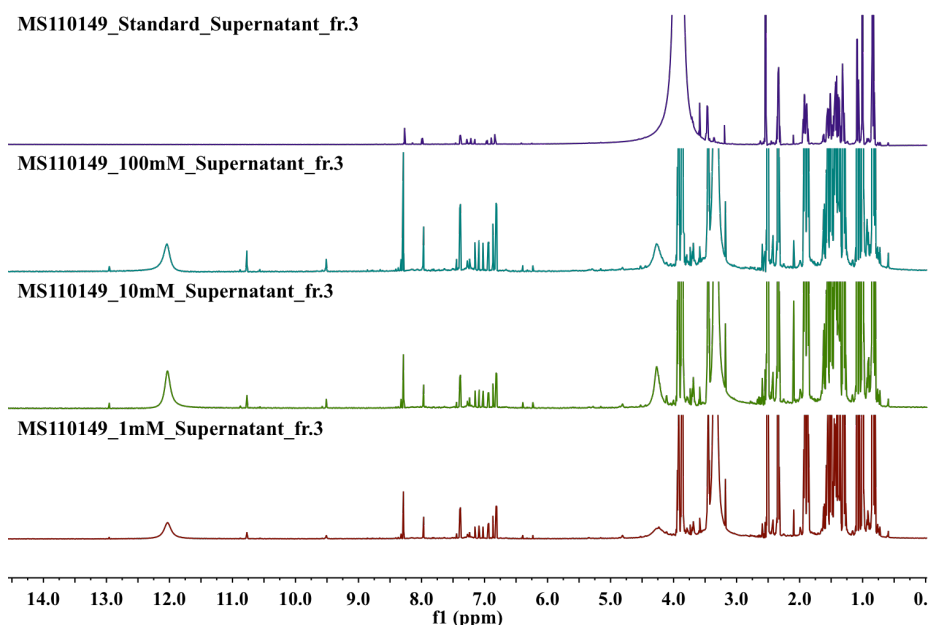
Figure 6.30 1H NMR spectra of the outliers of MS110149 indicated by PCA result of region 1. From bottom to top: MS110149_EtOH 1 mM_Supernatant_fr.3, MS110149_EtOH 10 mM_Supernatant_fr.3, MS110149_EtOH 100 mM_Supernatant_fr.3 and the corresponding standard fraction.

found to be beneficial to accumulation of metabolites. The third fractions of supernatant samples with the addition of different concentrations were all identified as outliers in score plot of MS110149 dataset (Figure 6.29). The NMR spectra in Figure 6.30 showed

there was no significant difference in signal intensities with 1 mM and 10 mM ethanol added. When ethanol level was reached to 100 mM, metabolites production was significantly enhanced. The effect of ethanol on metabolites biosynthesis was also confirmed by a series of studies and the mechanism was reported as multiple responses of intracellular signals, gene transcription and enzyme activity levels.[85, 172, 183]

The third type outliers were identified because of the different metabolites produced under specific OSMAC conditions and are thus the prioritized fractions in this project, especially those outliers identified from more than one NMR region (Figure 6.31). 11 outliers from marine strain MS110109 with their sample codes were shown in Figure 6.31 and 3 of them appeared in more than one score plot. After visual comparison with standard fractions, 6 fractions were determined to contain the "prioritized" outliers with different NMR signals (Figure 6.32), including the three fractions presumed to contain the same outlier.
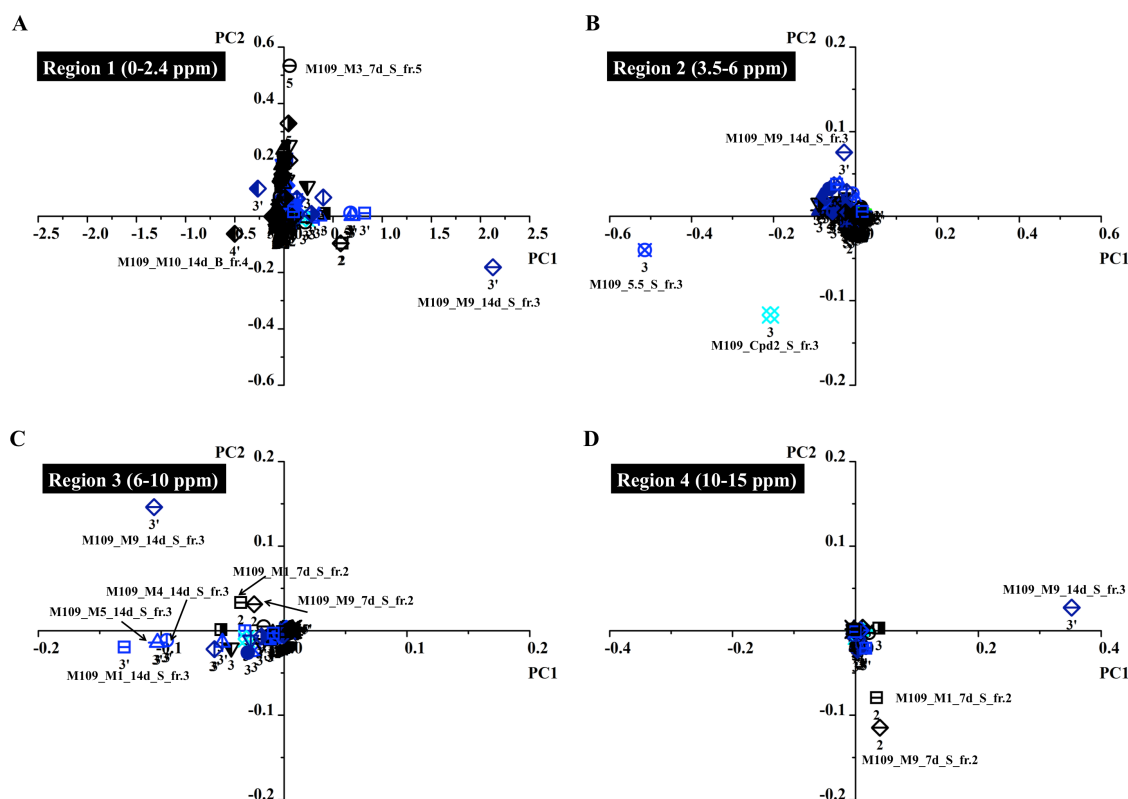
Figure 6.31 Score plots of 320 fractions from MS110109. Fraction numbers were shown below each symbol. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 – 2.4 ppm). B. Score plot of region 2 (3.5 – 6 ppm). C. Score plot of region 3 (6 – 10 ppm). D. Score plot of region 4 (10 – 15 ppm).
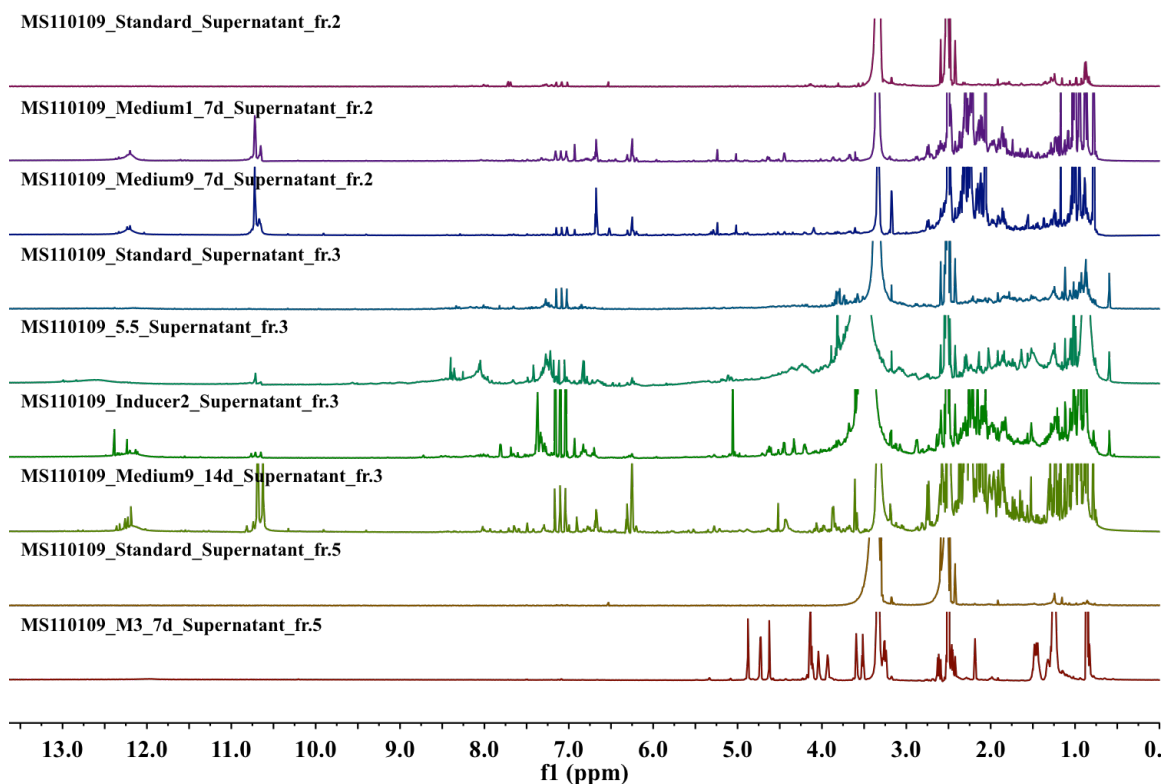
Figure 6.32 ^{1}H NMR spectra of the outliers of MS110109 and the corresponding standard fraction.

Visual inspection and comparisons of 156 outliers with their standard fraction in Table 6.4 confirmed the differences indicated by the corresponding loading plot. However, as mentioned before, some of the outliers containing strong signals from sample preparation process or the same signals with that in standard fractions were not selected as prioritized outliers in this project. As a result, 37 fractions with newly generated NMR signals under specific OSMAC conditions were prioritized for future investigation (Table 6.5). Among all the prioritized outliers, fractions derived from media variations (62.16%) gave more than that from stress conditions (37.84%) (Figure 6.33). Overall, cultivation for a longer time (14 days) provided more promising outliers than the standard culture period (7 days), containing 43.2% and 18.9 % of outliers, respectively. Within the 10 selected media, medium 1 (10.8 %) and 10 (10.8 %) were the most effective media in inducing unique metabolite production, whereas, none of the prioritized fractions were generated from medium 2, 5 or 8. Apart from media variation, stress shocks, including heat shock, pH shock and ethanol shock, as well as the addition of small inducer molecules also led to the production of unique metabolites. PH shock, especially acidic conditions (pH 3.5 and

5.5) provided 9 prioritized outliers, which are 64.3 % of the outliers induced by all stress conditions. The second effective strategy among stress conditions was found to be the use of inducers, with 1 outlier induced by inducer 1 and 2 outliers induced by inducer 2, followed by the high temperature shock condition (42 ℃), with 2 outliers.

Table 6.5 Prioritized outliers identified from PCA results of 13 actinomycetes

| Strains | Media variation | | Stress conditions |
	7 days	14 days	
ES120055	M10_7d_S_2	**M9_14d_S_3**	**3.5_S_3**, **5.5_S_3**
ES120127			**3.5_B_3**
LS120167		**M1_14d_S_3**, **M7_14d_S_3**	
LS120194		**M3_14d_S_3**, **M4_14d_S_3**, **M6_14d_S_3**	**42_S_3**
MS110104		**M4_14d_S_3**, **M10_14d_S_3**	**42_B_4**, **3.5_B_3**, **Inducer1_S_3**, **Inducre2_S_3**
MS110105	M10_7d_S_2	**M1_14d_S_3**, **M3_14d_S_3**, **M10_14d_S_3**	
MS110109	M1_7d_S_2, M3_7d_S_5, M9_7d_S_2	**M9_14d_S_3**	**5.5_S_3**, **Inducer2_S_3**
MS110115		M4_14d_B_4	**5.5_B_3**, **5.5_S_3**
MS110142		**M7_14d_S_3**	
MS110149	**M6_7d_S_3**		
MS110158			
MS110167			**3.5_S_3**, **5.5_S_3**
MS120045	M7_7d_S_3	M1_14d_S_3, **M6_14d_S_3**	
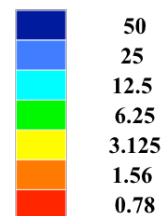
■	**50**
■	**25**
■	**12.5**
■	**6.25**
■	**3.125**
■	**1.56**
■	**0.78**

Anti-BCG activity was shown in different colors coded according to the legend (blue lowest and red highest activity)
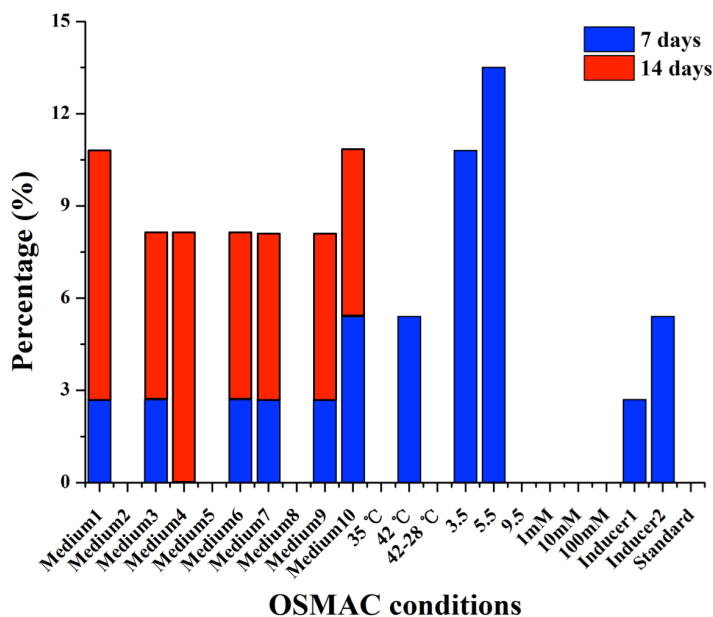
Figure 6.33 The distribution of 37 prioritized outliers.

The comparison of 37 outliers was explored by re-subjecting their NMR fingerprint data to PCA analysis. The resulting score plots of PC1 versus PC2 flagged up a few distinct

outliers in different NMR chemical shift regions (Figure 6.34). Under comprehensive consideration of PCA results and their anti-BCG activity, especially those active fractions with improved activity comparing to the corresponding standard fractions, 6 fractions derived from 5 strains were finally selected as the most potent fractions for further chemical investigation, including 3 fractions from pH 5.5 (ES120055_5.5_S_fr.3, MS110115_5.5_S_3 and MS110167_5.5_S_3), 1 fraction from medium3 (LS120194_Medium3_14d_S_3), 1 fraction from the addition of inducer 2 (MS110104_Inducer2_S_3) and 1 fraction from the high culture temperature condition (MS110104_42_B_4).
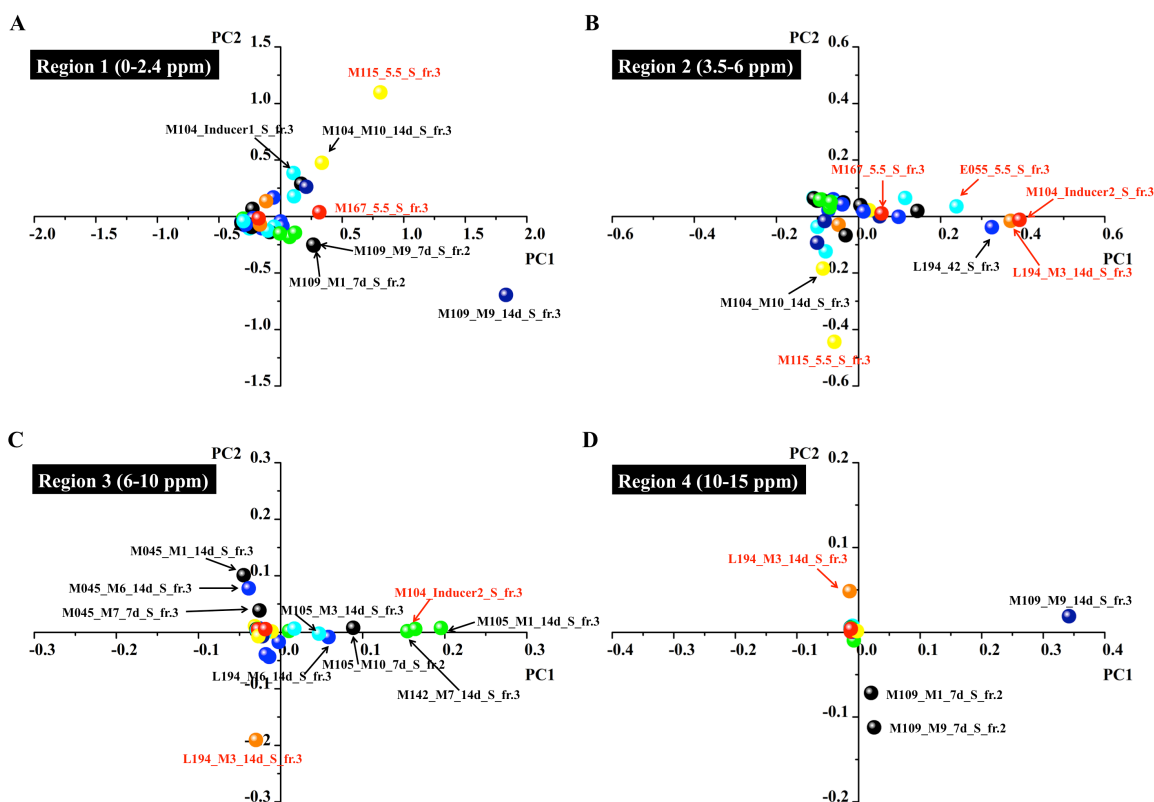
Figure 6.34 Score plots of 37 outliers fractions. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 – 2.4 ppm). B. Score plot of region 2 (3.5 – 6 ppm). C. Score plot of region 3 (6 – 10 ppm). D. Score plot of region 4 (10 – 15 ppm).

6.4 Conclusion

In an attempt to uncover the biosynthetic potential of 13 unique environments-derived actinomycetes, the effects of 32 OSMAC conditions, involving media variation and stress conditions, were investigated. A platform with the combination of HPLC fractionation

and NMR fingerprinting generated a huge data set of 4160 fractions. A whole cell-based screening has excluded 70% non-active samples and resulted in 1271 active candidate fractions (Figure 6.35). When discriminant analysis was applied on ^{1}H NMR spectra of active fractions, PCA produced a set of synthetic variables that clearly separated samples with newly generated metabolites grown under specific conditions with the others. It is a quick screen to expose outliers, although additional visual inspection of the original raw data (^{1}H NMR spectra) is required. As a result, 37 prioritized outliers were highlighted by PCA and 6 of them with high anti-BCG activity and unique NMR signals were considered as the most promising fractions for further investigation.
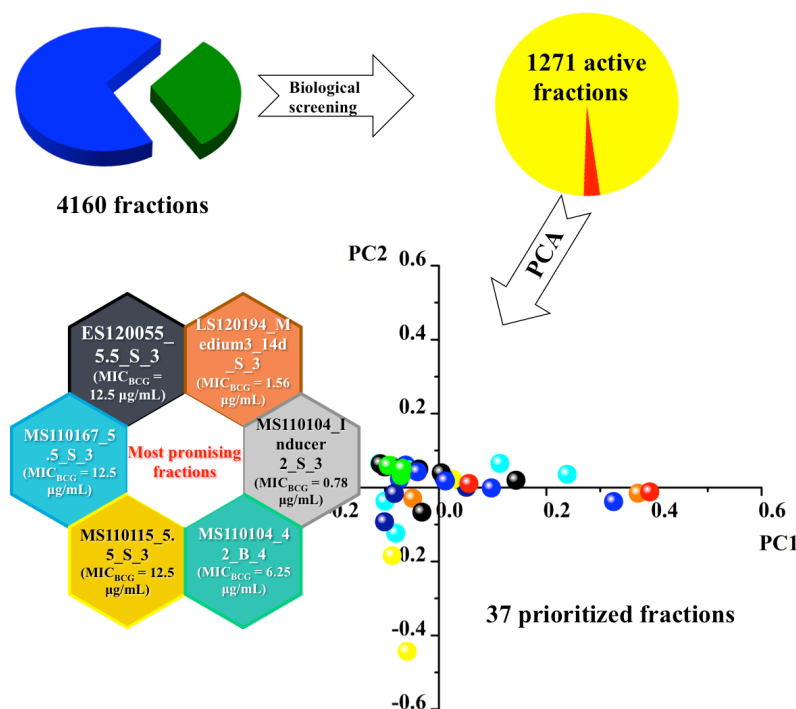
Figure 6.35 Process of highlighting promising fractions for further investigation.

Comparing to media variation, the use of stress conditions appeared to have stronger effects on triggering the silent biosynthetic genes of studied actinomycetes. Among the applied stress types, pH shock and the addition of inducer molecules displayed the biggest potential, whereas ethanol shock played a significant role in increasing the metabolites production rather than activating the expression of silent genes.

Chapter VII Target identification of anti-TB compounds from Hit fractions

Abstract: The three most active outliers identified in Chapter 6 were scaled up and 20 compounds were isolated, including one novel compound and 3 synthetic compounds that are reported for the first time as naturally occurring metabolites. Their structures were determined by means of HRESIMS and NMR data. Anti-microbial activity evaluation of 11 identified compounds to date resulted in 8 anti-BCG active compounds, with the most potent MIC value of 5 μg/mL. This is the first report that investigates the anti-BCG properties of the isolated compounds.

7.1 Introduction

Actinomycetes are a crucial source of novel natural products, but discovering more potential compounds that are biosynthesized by the so-called cryptic gene clusters represents a continuing challenge for natural product research. During the exploration of structure diversity by activating the cryptic pathway, the OSMAC approach has been proven to be convenient in maximizing the secondary metabolites of a single strain by varying the composition of media, changing condition of cultivation, adding precursors or enzyme inhibitors or co-cultivation.[84] In our study, 13 actinomycetes have been shown as the promising producers of anti-TB natural products (chapters 3 and 5). Guided by the OSMAC approach, 32 different conditions were attempted and the NMR profiles of the fractions were evaluated by PCA. As a result, 6 fractions of 5 actinomycetes with strong anti-BCG activity as well as novel NMR signals were selected as the most potential fractions containing unique and active compounds for further investigation (chapter 6). Scaled up fermentation and chemical exploration of three fractions with the strongest activity was performed, namely MS110104_Inducer2_Supernatant_fr.3, MS110167_ pH5.5_Supernatant_ fr.3 and LS120194_Medium3_14d_Supernatnat_fr.3. Details of the isolation, structure elucidation, and bioactivity screening of isolated compounds are reported in this chapter.

7.2 Large-scale isolation of anti-TB compounds from MS110104_Inducer2 _Supernatant_fr.3

Marine actinomycete MS110104 was isolated from a sediment sample collected from the South China Sea. This isolate was identified as a *Streptomyces* sp. based on cultural, physiological, morphological characteristics and 16s rRNA gene analysis (Figure S1).

In our search for bioactive secondary metabolites from unique environments-derived actinomycetes, MS110104 was found to be active against BCG (Chapter 3). With the aim to activate the expression of the silent biosynthetic gene clusters in this strain, a so-called OSMAC approach was carried out to find active compounds (Chapter 6). A total of 32 fermentation conditions were applied to MS110104 and further fractionation of the crude extracts generated 320 fractions. A cell-based anti-BCG screening of the fractions resulted in 153 active hits (47.8%) (Figure S2). Fraction 3 from the supernatant extract with the addition of the inducer 2 (MS110104_Inducer2_Supernatant_fr.3) is one of the

most active fractions.

Then the multivariate statistical analysis was conducted on the data set consisting of NMR fingerprint data of 320 fractions. The potential fraction MS110104_Inducer2_ Supernatant _fr.3 was highlighted as the outlier comparing to other fractions within chemical shifts from 0- 2.4 ppm and 3.5 -6 ppm regions (Figure S3). Visual inspection of the corresponding ^{1}H NMR spectra clearly showed a series of newly generated NMR signals only existed in MS110104_Inducer2_Supernatant_fr.3 (Figure 7.1).

Thus, MS110104 was scaled up (4 L) under the condition of additional inducer molecule 2 and the n-butanol extract was injected to the analytical C-18 column under the same gradient as described in Chapter 5 to generate the target fraction 3. With the use of Sephadex LH-20 column and semi-preparative HPLC, 5 compounds (**85 – 89**) were obtained from this target fraction.
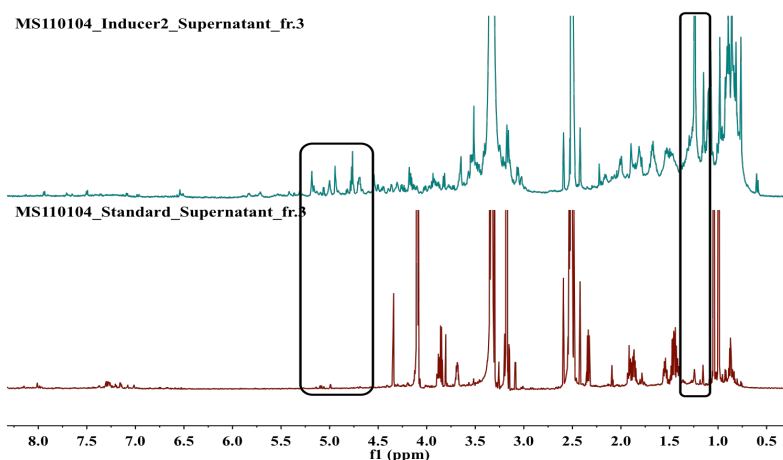
Figure 7.1 ^{1}H NMR spectra of the outlier MS110104_Inducer2_Supernatant_fr.3 indicated by PCA result of regions 1 and 2 and the corresponding standard fraction.

Compound **85** was isolated as a white powder. HRESIMS of **85** afforded an [M-H]$^-$ ion at m/z 621.4002 (calculated for 621.4008), consistent with a molecular formula of $C_{35}H_{58}O_9$. The ^{1}H NMR spectrum showed 58 proton signals including five olefinic protons at δ_H 6.53 (s), 5.83 (d, J=8.0 Hz), 5.71 (d, J= 11.4 Hz), 6.53 (dd, J= 15.7, 10.8 Hz) and 5.15 (dd, J= 14.4, 8.4) ppm, two methoxy protons at δ_H 3.54 (s) and 3.16 (s), nine methyl protons at δ_H 0.72 ~ 1.79 and six oxygenated methine protons at δ_H 3.17 ~ 5.13 ppm (Table 7.1). The ^{13}C NMR spectrum displayed 35 carbon signals that were classified by HSQC and HMBC spectra as two methoxy carbons at δ_C 59.8 and 55.6 ppm, nine methyl carbons at δ_C 7.4 ~ 23.1 ppm, five olefinic methine carbons at δ_C 132.3, 144.6, 124.3,

132.6 and 126.0 ppm, seven oxygenated methine carbons at δ_C 69.1~ 99.1 ppm, three fully substituted olefinic carbons (one oxygenated) and one conjugated ester carboxyl carbon at δ_C 164.9 ppm. Inspection of the 1D and 2D NMR obtained from **85** revealed a 16-membered macrolide (Table 7.1). The structure of the 16-membered macrolide was assigned on the basis of COSY spectrum and $^1H – {}^{13}C$ correlations in HMBC spectrum. The HMBC correlation of the proton at δ_H 5.13 (H-15) with the carbon at δ_C 164.9 (C-1), the proton at δ_H 6.53 (H-3) with the carbons at δ_C 164.9 (C-1), 141.3 (C-2), the proton at δ_H 5.83 (H-5) with the carbons at δ_C 132.3 (C-3) and 14.1 (C-26) and the proton at δ_H 5.71 (H-11) with the carbons at δ_C 41.7 (C-12) and 19.7 (C-29) indicated the linkage between C-1 and C-15 via oxygen, C-1 and C-3 via the fully substituted olefinic C-2, C-3 and C-5 via the fully substituted olefinic C-4 and C-9 and C-11 via the fully substituted olefinic C-10, which determined a 16-member lactone ring. The protons of two methoxy groups at δ_H 3.54 and 3.16 ppm were connected with the carbons at δ_C 141.3 (C-2) and 83.4 (C-14), respectively, based on their HMBC correlations. According to the correlations between the protons at δ_H 3.32 (H-23), 5.33 (19-OH), 0.88 (H-31) and the carbon at δ_C 99.1 ppm (Table 7.1), the pyranose fragment was formed and attached to C-18. Thus, the structure of **85** was dereplicated as the known compound bafilomycin A1 (Figure 7.2). The NMR data and optical rotation observed for **85** is consistent with those of bafilomycin A1, indicating that **85** could have the same absolute configuration.[184]

85: R$_1$=H, R$_2$=H
86: R$_1$=CH$_3$, R$_2$=H
87: R$_1$=CH$_3$, R$_2$=CH$_3$
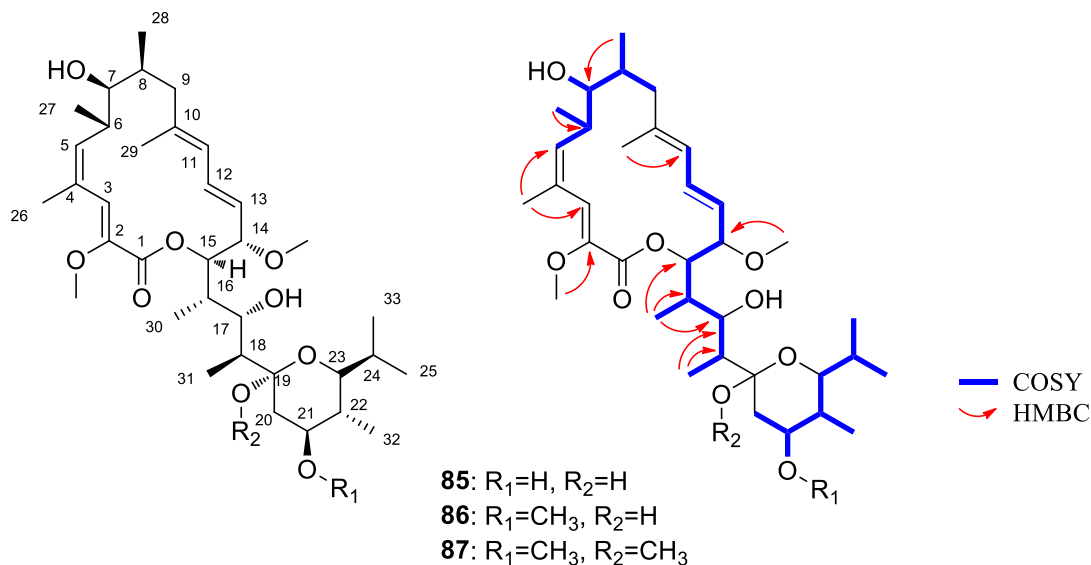
Figure 7.2 Structures and observed COSY and key HMBC correlations of **85** to **87**.

Table 7.1 ^{1}H and ^{13}C (800 MHz) NMR data of **85** to **87** in DMSO-d_6

Position	85		86		87		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
1	164.9 C		166.7 C		166.4 C			
2	141.3 C		141.8 C		141.6 C			
3	132.3 CH	6.53 s	133.2 CH	6.55 s	132.6 CH	6.55 s		1, 2, 5, 26
4	131.1 C		133.4 C		133.2 C			
5	144.6 CH	5.83 d (8)	141.0 CH	5.83 d (9.0)	142.1 CH	5.81 d (9.0)	6	3, 6, 7, 26
6	37.7 CH	2.42 m	37.2 CH	2.44 m	37.6 CH	2.44 m	5, 7, 27	4, 5, 27
7	78.9 CH	3.17 m	80.0 CH	3.20 m	80.4 CH	3.16 m	6, 8, 7-OH	5, 28
8	43.0 CH	1.6 dd (1.6, 7.2)	41.8 CH	1.65 m	42.6 CH	1.71 m	7, 9, 28	28
9	41.7 CH$_2$	1.88 m	40.6 CH$_2$	1.88 m	40.5 CH$_2$	1.87 m	8, 9'	10, 11, 29
		2.04 dd (11.2, 13.6)		2.10 m		2.12 m	8, 9	10, 11, 29
10	143.2 C		143.7 C		143.7 C			
11	124.3 CH	5.71 d (11.4)	125.5 CH	5.76 d (10.5)	128.1 CH	5.8 d (11.0)	12	9, 12, 13, 29
12	132.6 CH	6.53 dd (15.7, 10.8)	132.0 CH	6.64 dd (15.0, 11.0)	132.0 CH	6.63 dd (15.2, 10.6)	11, 13	11, 14
13	126.0 CH	5.15 dd (14.4, 8.4)	126.5 CH	5.21 dd (15.0, 9.5)	129.5 CH	5.23 dd (15.0, 9.4)	12, 14	11, 14, 15
14	83.4 CH	3.97 dd (7.2, 8)	83.0 CH	3.87 dd (8.5, 8)	83.0 CH	3.81 dd (8.2, 8)	13, 15	12, 13, 14-OCH$_3$, 15, 16
15	75.8 CH	5.13 dd (1.6, 6.4)	76.2 CH	5.05 d (9.0)	76.7 CH	5.05 dd (8.5, 1.7)	16	1, 14, 16, 17, 30
16	39.0 CH	1.86 m	38.8 CH	1.87 m	38.3 CH	1.91 m	15, 17, 30	17, 30
17	70.4 CH	3.99 m	70.9 CH	4.03 m	72.0 CH	4.00 m	16, 18, 17-OH	15, 16, 31
18	39.9 CH	1.72 m	41.3 CH	1.78 m	40.6 CH	1.77 m	17, 19, 31	16, 20
19	99.1 C		98.4 C		103.5 C			
20	43.3 CH$_2$	2.0 dd (4.8, 12.0)	43.3 CH$_2$	2.19 m	44.5 CH$_2$	2.08 m	19, 20', 21	19, 21, 22
		1.07 td (1.6, 12.0)		1.14 m		1.17 m	19, 20, 21	21, 22
21	69.1 CH	3.36 m	80.2 CH	2.94 m	77.8 CH	2.95 m	20, 22, 21-OH	19, 23, 32
22	41.2 CH	1.14 m	35.9 CH	1.48 m	36.2 CH	1.48 m	21, 23, 32	20, 21, 23
23	76.1 CH	3.32 m	71.3 CH	3.57 m	71.1 CH	3.57 m	22, 24	19, 21, 24, 25, 32
24	28.0 CH	1.82 td (2.4, 7.2)	28.1 CH	1.90 m	29.0 CH	1.92 m	23, 25, 33	25, 32
25	14.9 CH$_3$	0.72 d (7.2)	14.0 CH$_3$	0.80 d (6.7)	14.2 CH$_3$	0.80 d (6.7)	24	23, 24
26	14.1 CH$_3$	1.89 s	14.1 CH$_3$	2.00 s	14.1 CH$_3$	2.00 s		3, 4, 5
27	18.4 CH$_3$	0.96 d (7.2)	17.4 CH$_3$	0.98 d (7.0)	17.4 CH$_3$	0.96 d (7.2)	6	6
28	23.1 CH$_3$	0.89 d (6.5)	22.5 CH$_3$	0.95 d (6.5)	22.1 CH$_3$	0.94 d (6.5)	8	7
29	19.7 CH$_3$	1.79 s	20.2 CH$_3$	1.91 s	20.7 CH$_3$	1.93 s		11
30	11.0 CH$_3$	0.78 d ((7.2)	12.3 CH$_3$	0.87 d ((6.8)	12.2 CH$_3$	0.88 d ((6.8)	16	15, 16, 17
31	7.4 CH$_3$	0.88 d (3.2)	7.3 CH$_3$	1.00 d (7.2)	7.4 CH$_3$	0.98 d (7.2)	18	17, 18, 19
32	21.5 CH$_3$	0.87 d (6.4)	22.5 CH$_3$	0.84 d (7.5)	22.5 CH$_3$	0.84 d (7.4)	22	21, 23
33	12.8 CH$_3$	0.83 d (6.4)	14.3 CH$_3$	0.83 d (7.5)	14.2 CH$_3$	0.83 d (7.4)	24	23, 24, 25
2-OCH$_3$	59.8 CH$_3$	3.54 s	60.2 CH$_3$	3.56 s	60.8 CH$_3$	3.54 s		2
14-OCH$_3$	55.6 CH$_3$	3.16 s	55.8 CH$_3$	3.25 s	55.8 CH$_3$	3.22 s		14
19-OCH$_3$					49.6 CH$_3$	3.32 s		19
21-OCH$_3$			57.4 CH$_3$	3.30 s	57.8 CH$_3$	3.30 s		21

7-OH	4.91 d (5.6)	4.91 d (5.6)	4.90 d (5.8)	7	6, 7
17-OH	4.52 d (5.6)	4.54 d (6.2)	4.51 d (5.6)	17	16, 18
19-OH	5.33 d (1.6)	5.35 d (1.6)			19, 20
21-OH	4.46 d (5.6)			24	21, 22

86 gave a [M - H]$^-$ ion at m/z at 635.4161 (calculated for 631.4165) in the negative ion of HRESIMS, consistent with a molecular formula of $C_{36}H_{60}O_9$, differing from the molecular weight of bafilomycin A1 (**85**) simply by the addition of CH_2. Comparison of the 1D and 2D NMR data obtained for **86** (Table 7.1) with that we collected for bafilomycin A1 (**85**) showed that **86** was closely related to **85**. A HMBC correlation observed between a methyl ether resonance at δ_H 3.30 ppm and a carbon resonance at δ_C 80.2, assigned to C-21 in **86** by COSY and HSQC data, confirmed that the C-21 OH present in bafilomycin A1 (**86**) had been replaced by a methoxy in **86**. By comparison with the reported data, **86** was identified as bafilomycin G. [185]

87 gave a [M - H]$^-$ ion at m/z at 649.4317 (calculated for 649.4321) in the negative ion of HRESIMS, consistent with a molecular formula of $C_{37}H_{62}O_9$, differing from the molecular formula of bafilomycin G (**86**) simply by the addition of CH_2. The 1D and 2D NMR data obtained for **87** (Table 1) were nearly identical to the data obtained for bafilomycin G (**86**), indicating that the compounds were closely related. The major difference in the NMR data for **86** and **87** was that the OH-19 proton at δ_H 5.35 in **86** was replaced by a methyl singlet (19-OCH$_3$: δ_C 49.6; δ_H 3.32), suggesting that **87** was the methyl ketal of bafilomycin G (**87**). A correlation observed between the methyl resonance at δ_H 3.32 and the C-19 ketal carbon (δ_C 103.5) in the HMBC spectrum obtained for **87** confirmed this assignment. By comparison with the reported data, **87** was dereplicated as known compounds bafilomycin H.[185]

88 gave a [M - H]$^-$ ion at m/z at 603.3996 (calculated for 609.3902) in the negative ion of HRESIMS, consistent with a molecular formula of $C_{35}H_{56}O_8$, which differed from the molecular formula of bafilomycin A1 (**85**) by the loss of H_2O. Comparison of the 1D and 2D NMR data obtained for **88** (Table 7.2) with that of **85** to **87** (Table 7.1) showed that these compounds were closely related. In particular, the NMR data showed that the C-1 to C-15 macrocyclic rings were identical. Similarly, the NMR data confirmed that the substituents at C-22 (methyl), and C-24 (2 methyl groups) were also the same. However,

Table 7.2 ^{1}H and ^{13}C (800 MHz) NMR data of **88** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1	167.3 C			
2	141.4 C			
3	132.2 CH	6.53 s		1, 2, 5, 26
4	131.3 C			
5	143.1 CH	5.84 d (9.3)	6	3
6	36.9 CH	2.42 m	5, 7, 27	27
7	80.4 CH	3.17 dd (2.3, 6.8)	6, 8, 7-OH	9
8	43.3 CH	1.65 m	7, 9, 28	28
9	41.5 CH$_2$	1.89 m	8	11, 29
		2.05 m		
10	143.8 C			
11	125.4 CH	5.73 d (10.4)	12	12, 29
12	132.7 CH	6.54 dd (15.1, 10.4)	11, 13	11, 14
13	126.5 CH	5.17 m	12, 14	11, 14, 15
14	82.8 CH	3.98 m	13, 15	12, 13, 14-OCH$_3$, 15
15	75.9 CH	5.11 m	16	14
16	38.8 CH	2.07 m	15, 17, 30	31
17	72.1 CH	3.96 m	16, 18, 17-OH	15, 19, 30, 31
18	46.9 CH	2.99 m	17, 31	16, 20
19	202.6 C			
20	127.4 CH	6.20 d (15.67)	21	18, 22
21	147.7 CH	6.80 dd (8.75, 15.67)	20, 22	19, 23, 32
22	38.5 CH	2.53 m	21, 23, 32	20, 24
23	83.7 CH	3.19m	22, 24, 23-OH	21, 25, 32, 33
24	31.4 CH	1.90 m	23, 25, 33	22, 25, 33
25	17.8 CH$_3$	0.80 d (6.7)	24	23, 24, 33
26	14.1 CH$_3$	2.00 s		3, 5
27	18.4 CH$_3$	0.94 d (7.0)	6	6
28	23.5 CH$_3$	0.89 d (6.5)	8	7
29	20.2 CH$_3$	1.79 s		11
30	12.3 CH$_3$	0.96 d ((6.8)	16	15, 16, 17
31	9.5 CH$_3$	1.16 d (7.0)	18	17, 18, 19
32	15.6 CH$_3$	1.11 d (7.0)	22	21, 23
33	14.3 CH$_3$	0.91 d (7.5)	24	23, 24, 25
2-OCH$_3$	60.2 CH$_3$	3.55 s		2
14-OCH$_3$	55.8 CH$_3$	3.15 s		14
7-OH		4.89 d (6.56)	7	6, 8
17-OH		3.77 d (6.63)	17	16, 18
23-OH		3.23 d (6.63)	23	22, 24

unlike in **85**, the methine proton resonance at δ_H 2.53 (H-22) in **88** showed a COSY correlation to a olefinic methine resonance at δ_H 6.80 (H-21) and HMBC correlations to olefinic carbon resonances at δ_C 127.4 (C-20) and 31.4 (C-24), indicating the presence of an alkene. Comparing to **85**, the H-23 methine resonance at δ_H 3.19 in the ^{1}H NMR spectrum of **88** showed an additional COSY correlation to a hydroxy resonance at δ_H 3.23 (23-OH), suggesting the pyranose ring was open in **88**. [13]C spectrum of **88** displayed a downfield carbonyl signal at δ_C 202.6 ppm (C-19) and the HMBC correlations from protons at δ_H 3.96 (H-17), 6.80 (H-21) and 1.16 (H-31) to this carbonyl confirmed the linkage between C-20 and C-19 (Table 7.2). Comparison with reported data confirmed **88** as bafilomycin D (Figure 7.3).[186]
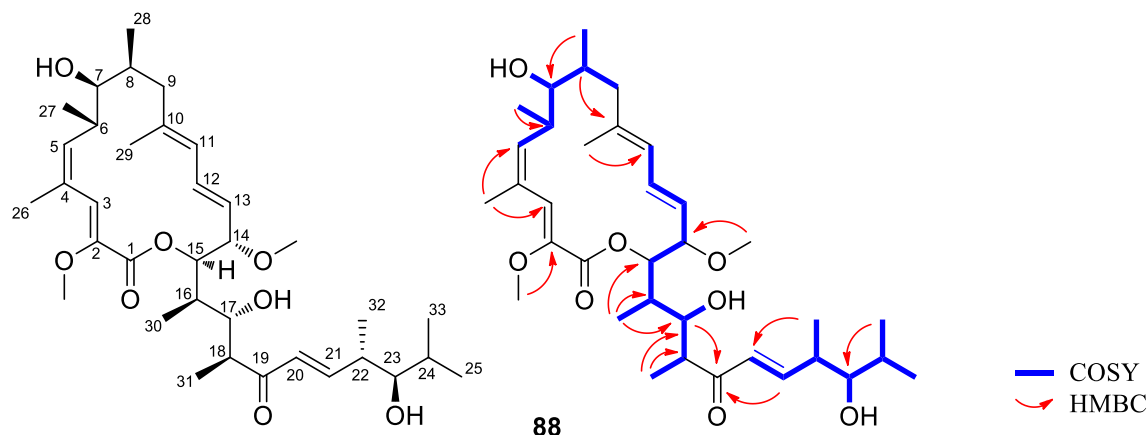
Figure 7.3 Structures and observed COSY and key HMBC correlations of **88**.

Antimicrobial assays of **85** − **88** against four Gram-positive bacteria and one Gram-negative bacterium were carried out using the method mentioned in Chapter 3 (Table 7.3). All of these bafilomycins showed the same level of activity against SA with MIC value of 100 μg/mL (62.2, 63.6, 65.0 and 60.4 μM, respectively). Compared to compounds **85** to **87**, **88** showed weaker anti-BCG activity with doubled MIC value of 40 μg/mL (24.16 μM), suggesting the presence of the pyranose moiety may play an essential role in the inhibition process.

To date, 19 bafilomycins have been isolated according to the Dictionary of Natural Products. These macrolides exhibit a variety of biological activities, including antitumor,[187-190] antibacterial,[186, 190] antifungal,[186, 190, 191] antiparasitic[192] and immunosuppressant activities.[193] It is worth mentioning that, among the bafilomycins identified in this work, there is no report in the literature about activity against *M. bovis*

BCG. In particular, bafilomycin A1 was found to be an extremely potent and specific inhibitor of the vacuolar ATPases.[194] The unique mode of mechanism drew considerable interest to the bafilomycins, especially in anti-TB drug discovery, which only has one candidate compound TMC 207 with the mechanism of ATPase inhibition.

Table 7.3 Antimicrobial activities of compounds **85 – 88**

MIC (μM)	Organism (Strain)				
	BCG	SA	MRSA	BS	PA
85	12.44	62.2	NA*	NA	NA
86	12.72	63.6	NA	NA	NA
87	13	65	NA	NA	NA
88	24.16	60.4	NA	NA	NA
Control	0.37[a]	0.7[b]	0.7[b]	0.35[b]	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

The comparison of NMR spectra between target fraction and isolated compounds is shown in Figure 7.4. All of the signals from this fraction could be explained by the pure compounds.
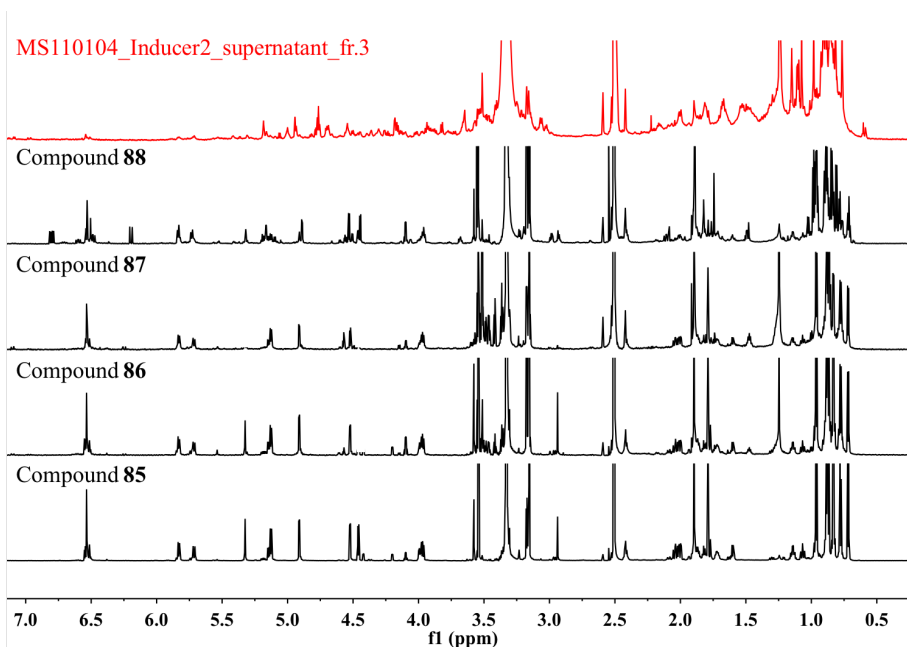
Figure 7.4 Comparison of NMR fingerprint between target fraction and compounds of MS110104.

7.3 Large-scale isolation of anti-TB compounds from MS110167_pH5.5 _Supernatant_fr.3

Marine actinomycete MS110167 was isolated from a sediment sample collected from the South China Sea. This isolate was identified as a *Streptomyces* sp. based on cultural,

physiological, morphological characteristics and 16s rRNA gene analysis (Figure S4). MS110167 was shown to have anti-BCG activity in a cell-based screening assay (chapter 3). The application of the OSMAC approach to this strain using 32 different culture conditions resulted in 320 fractions, 92 of which (28.75%) displayed activity against BCG in a cell-based screening assay (Chapter 6). The third fraction of supernatant extract generated from MS110167 under the medium pH at 5.5 (MS110167_pH5.5 _Supernatant_fr.3) is the most active fraction with MIC values of 0.78 μg/mL (Figure S5). A comparison between 320 fractions generated from different fermentation conditions was carried out with the use of PCA strategy. The most active fraction MS110167_pH5.5_Supernatant _fr.3 was identified as the one of the distinct outliers comparing to other fractions within chemical shifts from 0- 2.4 ppm and 3.5 -6 ppm regions (Figure S6). Visual inspection of the corresponding ^{1}H NMR spectra clearly depicted a dramatic shift in the profile of the secondary metabolites obtained compare to those obtained under the initial condition (Figure 7.5).
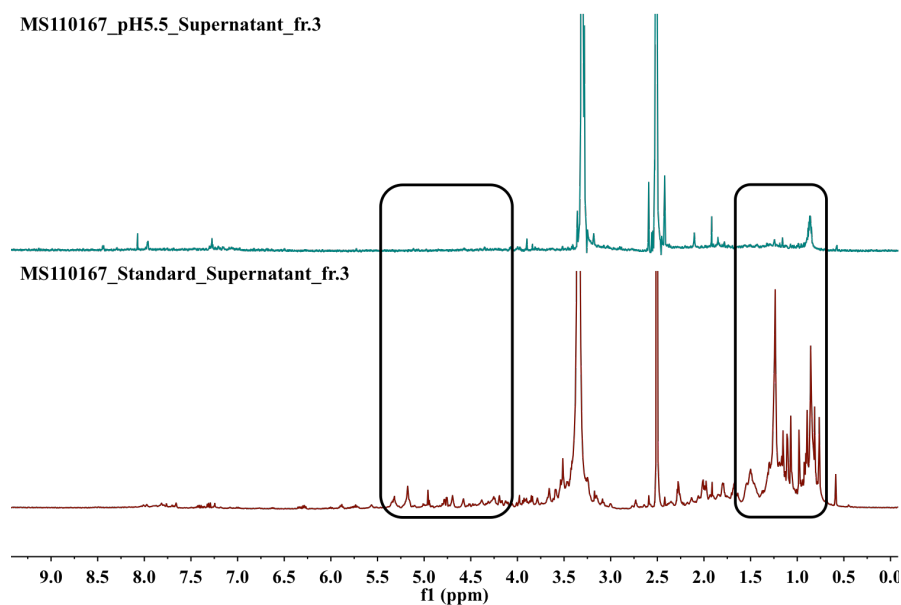
Figure 7.5 ^{1}H NMR spectra of the outlier MS110167_pH5.5_Supernatant_fr.3 indicated by PCA result of regions 1 and 2 and the corresponding standard fraction.

In order to obtain enough quantities of the newly produced compounds and characterize them, a 4 L large scale of MS110167 under the condition of pH 5.5 was conducted. The investigation resulted in the isolation of seven natural products (**89 – 95**).

89 was isolated as a colorless powder from the target fraction by Sephadex LH-20

followed by semi-preparative HPLC. Compound **89** showed UV absorption maximum at 320 nm, suggesting the presence of a conjugated system. The molecular weight of **89** was determined by HRESIMS with a [M+H]$^+$ ion at m/z 513.2935 (calculated for 513.2959), suggesting a molecular formula of $C_{29}H_{40}N_2O_6$. The ^{1}H NMR spectrum displayed 40 proton signals including four olefinic protons at δ_H 5.71 (dd, J= 2.0, 11.2 Hz), 5.91 (td, J=2.0, 11.0 Hz), 6.58 (dd, J= 10.0, 15.0 Hz) and 6.86 (d, J= 16.2 Hz) ppm, two methyl protons at δ_H 0.86 (t, J=7.2) and 1.06 (d, J=6.5 Hz), six methylene groups at δ_H 0.87 ~ 3.53 ppm, two oxygenated methine protons at δ_H 3.34 and 3.81 ppm and 4 exchangeable protons at δ_H 8.90 (s), 7.96 (t, J= 6.0 Hz), 4.47 (s) and 5.10 (s) ppm (Table 7.4). The ^{13}C NMR spectrum displayed signals for 29 carbons that were classified by HSQC and HMBC spectra as four olefinic methine carbons at δ_C 124.6, 139.5, 150.5 and 121.6 ppm, two methyl carbons at δ_C 13.3 and 19.0 ppm, six methylene carbons at δ_C 26.4 ~ 42.4 ppm, two oxygenated methine carbons at δ_C 75.2 and 70.5 ppm, one fully substituted olefinic carbon at δ_C 103.0 and four down-field carbonyl carbons at δ_C 165.7, 172.6, 176.3 and 193.5 ppm (Table 7.4). Two ^{1}H-^{1}H coupling networks were apparent in the COSY correlations, and the larger network on the lipophilic subunit was hence identified to have the 5,5,6- tricyclic skeleton containing an oxygenated methine carbon at δ_C 75.2 ppm. The other ^{1}H-^{1}H coupling network was identified as -CH-CH-CH$_2$-CH$_2$- according to their correlations in COSY spectrum. The amide proton at δ_H 7.96 ppm, one of the protons at the terminal methylene of the second COSY coupling system (3.26 ppm) and the olefinic protons at δ_H 5.71 and 5.91 ppm show a HMBC correlation to the same amide carbonyl carbon at δ_C 165.7 ppm, thus establishing the connectivity between the two units via the linkage of the amide carbonyl group. The other pair of olefinic protons also showed a HMBC correlation to a carbonyl carbon at δ_C 172.6 ppm as well as to a carbon at δ_C 103.0 ppm which in turn correlated to the amide proton at δ_H 8.90 ppm, proving the connectivity between the carbonyl carbon (172.6 ppm) and the fully substituted olefinic carbon (103.0 ppm). The HMBC correlations from the amide proton at δ_H 8.90 ppm to carbon at δ_C 193.5 and from the terminal methine proton of the second COSY coupling system (3.86 ppm) to the carbonyl carbon at δ_C 176.3 confirmed the structure of the tetramic acid moiety. Additionally, with the comparison with the reported data, **89** was assigned as dihydromaltophilin (Figure 7.6).[195] The coupling constants of the two pair

olefinic protons are typical for a Z-configuration in the case of 5.71 (H-2) and 5.91 ppm (H-3) with a *J* value of 11.0 Hz and an E-configuration for 6.58 (H-17) and 6.86 ppm (H-18) with a *J* value of 15.0 Hz.

89: R₁=H, R₂=OH, R₃=OH
90: R₁=H, R₂=OH, R₃=H
91: R₁=CH₃, R₂=OH, R₃=H
92: R₁=H, R₂=O, R₃=OH
93: R₁=H, R₂=O, R₃=H
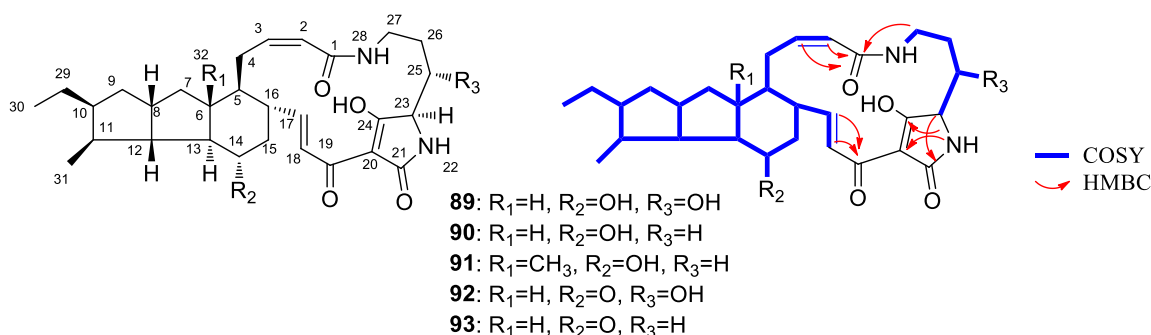
Figure 7.6 Structures and observed COSY and key HMBC correlations of **89 – 93**.

90 was obtained as colorless powder. Initial HRESIMS analyses gave a $[M+H]^+$ ion at m/z 497.3014 (calculated for 497.3010), indicating the molecular formula of $C_{29}H_{40}N_2O_5$, differing from the molecular formula of dihydromaltophilin (**89**) by the loss of oxygen. The ^1H and ^{13}C NMR signals indicated structure of **90** was very similar to that of **89**. The only observed difference between **89** and **90** was at the C-25 position, and the oxygenated methine in **89** was replaced by a methylene in **90**, indicating the loss of a hydroxyl group at C-25 (Table 7.4). Therefore, **90** was dereplicated to be a known compound 3-deOH-HSAF (Figure 7.6).[196]

91 was obtained as colorless powder and its molecular weight was determined by HRESIMS with a $[M+H]^+$ ion at m/z 511.3172 (calculated for 511.3166), consistent with the molecular formula of $C_{30}H_{42}N_2O_5$, which was differing from the molecular weight of 3-deOH-HSAF (**90**) simply by the addition of CH_2. The 1D and 2D NMR data obtained for **91** (Table 1) were nearly identical to the data obtained for **90**, except that the methine proton (H-6) at δ_H 1.67 ppm in **90** was replaced by a methyl singlet (32-CH₃: δ_C 16.0 ppm; δ_H 1.04 ppm). A correlation observed between the methyl resonance at δ_H 1.04 ppm and the carbons at δ_C 28.2 (C-4), 48.3 (C-5) and 37.6 ppm (C-6) in the HMBC spectrum obtained for **91** confirmed this assignment (Table 7.4). By comparison with the reported data, **91** was dereplicated as known compound ikarugamycin (Figure 7.6).[197]

92 was obtained as colorless powder and its molecular weight was determined by HRESIMS with a $[M+H]^+$ ion at m/z 511.2809 (calculated for 511.2803), differing from the molecular weight of dihydromaltophilin (**89**) simply by the addition of an unsaturated

Table 7.4 ^{1}H and ^{13}C (800 MHz) NMR data of **89 – 93** in DMSO-d_6

Position	89		90		91		92		93		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
1	165.7 C		168.7 C		165.1 C		167.0 C		165.1 C			
2	124.6 CH	5.71 dd (2.0, 11.2)	124.6 CH	5.78 dd (2.2, 11.0)	125.4 CH	5.74 dd (2.2, 11.6)	125.5 CH	5.75 dd (2.0, 11.0)	123.9 CH	5.71 dd (2.0, 11.2)	3	1, 4
3	139.5 CH	5.91 td (2.0, 11.0)	140.0 CH	5.94 td (2.2, 11.2)	137.9 CH	5.96 td (2.0, 10.8)	138.6 CH	5.89 td (2.0, 11.4)	137.9 CH	5.92 td (2.2, 11.2)	2, 4, 4'	1
4	28.5 CH₂	3.53 m	29.5 CH₂	3.44 m	28.2 CH₂	3.33 m	28.5 CH₂	3.50 m	27.3 CH₂	3.62 m	3, 4', 5	2
		1.90 dd (2.1, 17.0)		1.90 dd (2.1, 17.0)		1.94 m		1.90 m		1.98 m	3, 4, 5	2
5	44.1 CH	1.28 m	45.4 CH	1.30 m	48.3 CH	1.35 m	44.3 CH	1.24 m	42.8 CH	1.29 m	4, 4', 6, 16	4, 6, 7, 13, 15
6	48.1 CH	1.64 m	48.3 CH	1.67 m	37.6 C		49.0 CH	1.63 m	50.7 CH	1.68 m	5, 7, 13	4, 7, 13
7	37.9 CH₂	1.97 m	38.7 CH₂	2.01 m	49.5 CH₂	1.51 m	37.8 CH₂	2.00 m	38.0 CH₂	2.05 m	6, 7', 8	6, 9
		0.87 m		0.90 m		1.26 m		0.90 m		0.85 m	6, 7, 8	6, 9
8	43.3 CH	2.43 m	43.3 CH	2.43 m	37.1 CH	1.49 m	43.1 CH	2.46 m	43.1 CH	2.33 m	7, 9, 12	
9	40.7 CH₂	2.02 m	41.7 CH₂	2.07 m	39.1 CH₂	2.12 m	40.4 CH₂	1.95 m	39.4 CH₂	2.03 m	8, 9', 10	7, 12, 29
		0.81 m		0.85 m		0.88 m		0.80 m		0.82 m	8, 9, 10	7, 12, 29
10	54.1 CH	1.33 m	55.4 CH	1.30 m	54.5 CH	1.39 m	53.8 CH	1.27 m	52.9 CH	1.35 m	9, 11, 29	9, 29, 30, 31
11	46.8 CH	1.27 m	47.1 CH	1.30 m	47.1 CH	1.26 m	47.4 CH	1.24 m	46.4 CH	1.17 m	10, 12, 31	9, 13, 31
12	58.7 CH	1.76 m	59.0 CH	1.75 m	58.9 CH	1.77 m	50.9 CH	2.41 m	50.0 CH	2.20 m	8, 11, 13	7, 8, 9, 13, 14
13	59.6 CH	1.1 m	60.2 CH	1.19 m	60.0 CH	1.28 m	63.9 CH	2.10 m	62.6 CH	2.35 m	6, 12, 14	6, 7, 12, 15
14	75.2 CH	3.34 m	75.2 CH	3.34 m	74.0 CH	2.36 m	207.6 C		207.0 C		13, 15, 14-OH	6, 12, 13, 15
15	42.4 CH₂	1.75 m	43.3 CH₂	1.80 m	43.6 CH₂	1.72 m	47.2 CH₂	2.25 m	45.2 CH₂	2.27 m	14, 15', 16	
		1.26 m		1.3 m		1.27 m		2.06 m		2.10 m	14, 15, 16	
16	46.3 CH	2.05 m	46.7 CH	2.10 m	46.5 CH	2.11 m	47.4 CH	2.23 m	47.4 CH	2.27 m	5, 15, 15', 17	4, 6, 15
17	150.5 CH	6.58 dd (10.0, 15.0)	148.9 CH	6.52 dd (9.60, 15.4)	149.5 CH	6.68 dd (9.6, 15.0)	150.4 CH	6.67 dd (10.4, 15.2)	147.3 CH	6.65 dd (10.4, 15.4)	16, 18	15, 19
18	121.6 CH	6.86 d (16.2)	122.2 CH	6.91 d (16.0)	122.1 CH	6.90 d (16.2)	121.9 CH	6.84 d (16.4)	121.4 CH	6.94 d (16.0)	17	16, 19, 20
19	172.6 C		172.0 C		174.0 C		172.9 C		171.5 C			

Position	δC (1)	δH (1)	δC (2)	δH (2)	δC (3)	δH (3)	δC (4)	δH (4)	δC (5)	δH (5)	COSY	HMBC
20	103.0 C		103.0 C		103.1 C		102.9 C		103.4 C			
21	176.3 C		175.8 C		177.0 C		178.1 C		175.3 C			
22		8.90 s		8.92 s		8.95 s		8.95 s		8.80 s	23	20, 24
23	69.1 CH	3.86 m	70.0 CH	3.88 m	67.3 CH	3.87 m	69.1 CH	3.82 m	68.2 CH	3.86 m	22, 25	20, 21, 24, 26
24	193.5 C		195.2 C		196.6 C		194.5 C		192.6 C			
25	70.5 CH	3.81 m	28.3 CH2	2.24 m	27.0 CH	2.28 m	71.6 CH	3.86 m	28.7 CH2	2.20 m	23, 26, 26', 25-OH	23, 27
				2.06 m		2.03 m				2.01 m		
26	31.6 CH2	1.4 t (13.6)	27.7 CH2	1.40 m	27.7 CH2	1.51 m	32.3 CH2	1.42 m	27.6 CH2	1.36 m	25, 26', 27, 27'	25, 27
		1.2 m		1.24 m		1.20 m		1.25 m		1.18 m	25, 26, 27, 27'	25, 27
27	36.7 CH2	3.26 td (4.0, 9.2)	38.0 CH2	3.48 m	38.4 CH2	3.38 m	37.6 CH2	3.21 m	36.0 CH2	3.24 m	26, 26', 27', 28	1, 25, 26
		2.58 m		2.78 m		2.75 m		2.55 m		2.56 m	26, 26', 27, 28	25, 26
28		7.96 t (6.0)		7.95 t (6.0)		8.00 t (6.0)		8.00 t (6.0)		7.86 t (6.0)	27, 27'	1, 2, 26
29	26.4 CH2	1.56 m	27.4 CH2	1.58 m	27.2 CH2	1.55 m	26.1 CH2	1.50 m	25.1 CH2	1.50 m	10, 29', 30	30
		1.04 m		1.07 m		1.05 m		0.95 m		1.00 m	10, 29, 30	30
30	13.3 CH3	0.86 t (7.2)	13.2 CH3	0.88 t (7.2)	12.3 CH3	0.90 t (7.2)	12.7 CH3	0.81 t (7.2)	13.0 CH3	0.83 t (7.2)	29, 29'	10, 29
31	19.0 CH3	1.06 d (6.5)	19.0 CH3	1.08 d (6.6)	19.6 CH3	1.06 d (6.5)	18.9 CH3	1.07 d (6.5)	19.3 CH3	1.06 d (6.5)	11	10, 12
32					16.0 CH3	1.04 s						5, 6, 7
14-OH		4.47 s		4.49 s		4.49 s					14	13, 15
25-OH		5.10 s						5.08 s			25	23, 26

degree. The major difference in the NMR data for **89** and **92** was that the hydroxyl-attached methine (14-CH: δ_C 75.2 ppm; δ_H 3.34 ppm) in **89** was replaced by a non-conjugated carbonyl group at δ_C 207.6 ppm. HMBC correlations from protons at 1.63 (H-6), 2.41 (H-12) and 2.23 ppm (H-16) to the carbonyl carbon further confirmed the assignments (Table 7.4). With the supporting of published data, **92** was identified as compound xanthobaccin A (Figure 7.6).[198]

93 was obtained as colorless powder. HRESIMS data with a [M+H]$^+$ ion at 495.2859 (calculated for 495.2853), indicating that **93** had a molecular formula of $C_{29}H_{38}N_2O_5$, differing from that of xanthobaccin A (**92**) by the loss of oxygen. The ^{1}H and ^{13}C NMR data in combination with 2D correlations suggested **93** is structurally similar to **92**, except for the absence of a hydroxyl group at C-25 in **93** (Table 7.4). The structure of **93** was also confirmed by the COSY correlations between H-23, H-25 and H-26 as well as HMBC correlations from H-22 and H-27 to C-25 (Table 7.4). By comparison to the published data, **93** was dereplicated as xanthobaccin B (Figure 7.6).[199]

Apart from 5 polycyclic tetramate macrolactams (**89** – **93**), two triterpene glycosides were also identified from the same target fraction (Figure 7.7). **94** was obtained as a white powder and its molecular formula was determined as $C_{48}H_{78}O_{18}$ according to HRESIMS result ([M+Na]$^+$ 965.5086, calculated for 965.5080). The ^{1}H NMR displayed 78 protons, including an olefinic proton at δ_H 5.17 (t, J=3.2 Hz), 8 methyl groups from 0.76 to 1.11 ppm (7 singlet and one doublet), 11 methylenes from 0.94 to 3.93 ppm and three down-field oxygenated methine protons at δ_H 4.30 (d, J=8.0), 4.78 (d, J=7.2) and 4.96 ppm (d, J=7.6). ^{13}C NMR showed the presence of a carboxyl carbon signal at δ_C 172.3 ppm. A database search was conducted by using the above information and a known compound soyasaponin I was found to show high similarity with **94**.[200] The key HMBC correlations from proton H-3 to Glc-C1, Glc-H2 to Gal-C1 and Gal-H2 to Rha-C1 further confirmed the structure and the connectivity between triterpene skeleton and the sugar moiety (Table 7.5). Thus, **94** was determined to be soyasaponin I.

95 was isolated as a white powder and its molecular formula was determined as $C_{49}H_{80}O_{18}$ by HRESIMS result ([M+Na]$^+$ 979.5245, calculated for 979.5237), differing from the molecular weight of soyasaponin I (**94**) simply by the addition of CH_2. The ^{1}H and ^{13}C NMR spectra of **95** were in good agreement with those of **94**, except for the

presence of an additional methoxyl group in **94** (δ_C 52.1 ppm; δ_H 3.18 ppm). The HMBC correlation from the methoxyl proton to carbon at δ_C 170.3 ppm (Glc-C6) confirmed the structure of **95** as the methyl ester of **94**.[201]

94: R=H
95: R=CH₃
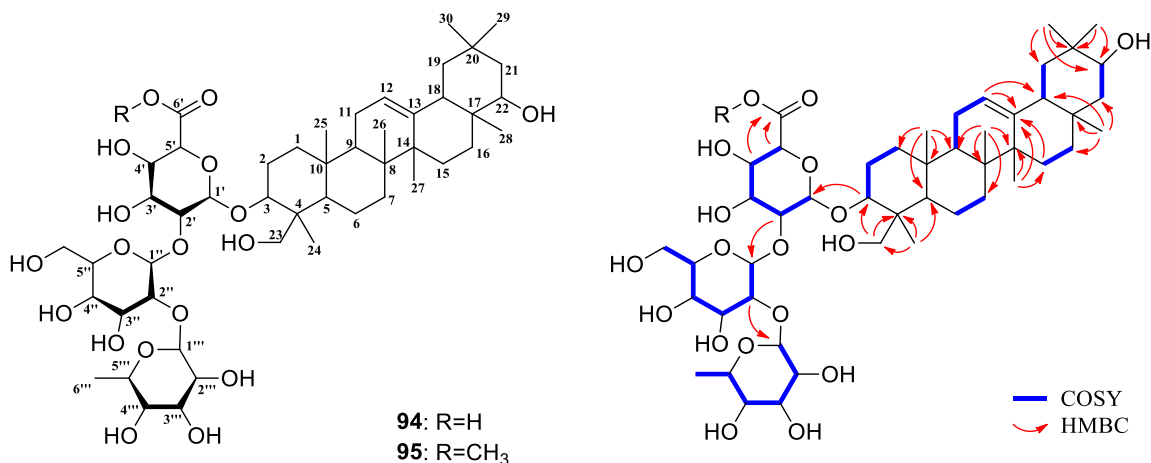
Figure 7.7 Structures and observed COSY and key HMBC correlations of **94** and **95**.

Table 7.5 ^{1}H and ^{13}C (800 MHz) NMR data of **94** and **95** in DMSO-d_6

Position	94		95		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
1	38.6 CH₂	1.62 m	39.2 CH₂	1.61 m	1', 2, 2'	3, 5, 25
		1.04 m		1.04 m	1,2, 2'	3, 5, 25
2	26.4 CH₂	2.13 m	26.6 CH₂	2.12 m	1, 1', 2', 3	1, 4
		1.84 m		1.83 m	1,1', 2, 3	1, 4
3	91.2 CH	3.27 m	90.9 CH	3.25 m	2, 2'	1, 23, 24, Glc-1
4	43.9 C		44.5 C			
5	56.1 CH	0.98 m	56.0 CH	0.97 m	6, 6'	1, 3, 7, 9, 23, 24, 25
6	18.5 CH₂	1.65 m	17.7 CH₂	1.66 m	5, 5', 7, 7'	4, 8, 10
		1.40 m		1.41 m	5, 5', 6, 7,7'	4, 8, 10
7	33.2 CH₂	1.54 m	33.1 CH₂	1.55 m	6, 6', 7'	5, 9, 14, 26
		1.29 m		1.29 m	6, 6', 7	5, 14, 26
8	39.9 C		40.0 C			
9	47.8 CH	1.55 m	47.9 CH	1.54 m	11	1, 5, 7, 12, 14, 25, 26
10	36.4 C		36.5 C			
11	24.0 CH₂	1.79 m	24.1 CH₂	1.77 m	9, 12	8, 10, 13
12	122.1 CH	5.17 t (3.2)	122.9 CH	5.15 m	11	9, 13, 14, 18
13	144.8 C		144.8 C			
14	42.3 C		42.4 C			
15	26.0 CH₂	1.67 m	26.7 CH₂	1.66 m	15', 16, 16'	8, 13, 17, 27
		0.95 m		0.96 m	15, 16, 16'	8, 13, 17, 27
16	28.7 CH₂	1.66 m	28.6 CH₂	1.67 m	15, 15', 16'	14, 17, 18, 28
		0.94 m		0.94 m	15, 15', 16	14, 18, 28

	δC	δH (J Hz)	δC	δH (J Hz)	COSY	HMBC
17	37.9 C		38.0 C			
18	45.3 CH	2.00 m	45.3 CH	2.02 m	19, 19'	12, 14, 16, 20, 22, 28
19	46.7 CH$_2$	1.93 m	46.8 CH$_2$	1.94 m	18, 19'	13, 17, 21, 29, 30
		1.44 m		1.46 m	18, 19	13, 17, 21, 29, 30
20	30.8 C		30.9 C			
21	42.2 CH$_2$	1.30 m	42.4 CH$_2$	1.28 m	21', 22	17, 19, 29, 30
		1.22 m		1.22 m	21, 22	19, 29, 30
22	75.5 CH	3.56 m	75.6 CH	3.56 m	21, 21'	16, 18, 20
23	23.0 CH$_3$	1.15 s	23.0 CH$_3$	1.15 s		3, 4, 5, 24
24	63.6 CH$_2$	3.93 d (10.4)	63.6 CH$_2$	3.93 d (10.4)	24'	3, 5, 23
		3.09 t (11.2)		3.09 t (11.2)	24	3, 5, 23
25	15.8 CH$_3$	0.81 s	15.8 CH$_3$	0.81 s		1, 5, 9, 10
26	17.0 CH$_3$	0.89 s	17.0 CH$_3$	0.89 s		7, 8, 9, 14
27	25.7 CH$_3$	1.07 s	25.7 CH$_3$	1.07 s		8, 13, 14, 15
28	21.1 CH$_3$	0.76 s	21.7 CH$_3$	0.76 s		16, 17, 18, 22
29	19.2 CH$_3$	0.86 s	19.6 CH$_3$	0.86 s		19, 20, 21, 30
30	28.6 CH$_3$	0.98 s	28.2 CH$_3$	0.98 s		19, 20, 21, 29
Glc-1	105.4 CH	4.30 d (8.0)	105.4 CH	4.30 d (8.0)	Glc-2	3, Glc-3, Glc-5
Glc-2	78.3 CH	3.66 m	78.7 CH	3.67 m	Glc-1, Glc-3	Glc-4, Gal-1
Glc-3	76.6 CH	3.56 m	76.7 CH	3.56 m	Glc-2, Glc-4	Glc-1, Glc-5
Glc-4	73.7 CH	3.24 m	73.4 CH	3.22 m	Glc-3, Glc-5	Glc-2, Glc-6
Glc-5	77.6 CH	3.48 m	77.4 CH	3.51 m	Glc-4	Glc-1, Glc-3, Glc-6
Glc-6	172.3 C		170.3 C			
Glc-7			52.1 CH$_3$	3.18 s		Glc-6, Glc-7
Gal-1	101.7 CH	4.78 d (7.2)	101.7 CH	4.77 d (7.2)	Gal-2	Glc-2, Gal-2, Gal-5
Gal-2	77.7 CH	3.50 m	76.6 CH	3.48 m	Gal-1, Gal-3	Gal-4, Rha-1
Gal-3	76.4 CH	3.55 m	76.5 CH	3.55 m	Gal-2, Gal-4	Gal-1, Gal-5
Gal-4	71.1 CH	3.24 m	71.2 CH	3.21 m	Gal-3, Gal-5	Gal-2, Gal-6
Gal-5	76.5 CH	3.52 m	77.0 CH	3.51 m	Gal-4, Gal-6, Gal-6'	Gal-1, Gal-5
Gal-6	61.5 CH$_2$	3.79 m	61.8 CH$_2$	3.77 m	Gal-5, Gal-6'	Gal-4, Gal-5
		3.65 m		3.67 m	Gal-5, Gal-6	Gal-4
Rha-1	102.3 CH	4.96 d (7.6)	102.2 CH	4.95 d (7.4)	Rha-2	Gal-2, Rha-3, Rha-5
Rha-2	72.3 CH	3.60 m	72.4 CH	3.61 m	Rha-1, Rha-3	Rha-4
Rha-3	72.7 CH	3.45 m	72.8 CH	3.45 m	Rha-2, Rha-4	Rha-1, Rha-3
Rha-4	74.3 CH	3.74 m	74.4 CH	3.75 m	Rha-3, Rha-5	Rha-2, Rha-6
Rha-5	69.3 CH	3.90 dd (6.4, 6)	69.4 CH	3.91 dd (6.4, 6.2)	Rha-4, Rha-6	Rha-1, Rha-4
Rha-6	18.9 CH$_3$	1.11 d (6.4)	18.7 CH$_3$	1.11 d (6.4)	Rha-5	Rha-4, Rha-5

Compounds **89** – **95** were evaluated for the growth inhibitory activity against microbial pathogens (Table 7.6). As shown in Table 7.6, only compounds **89** – **92** exhibited anti-BCG activity, whereas the other compounds were inactive. Among all the polycyclic

tetramate macrolactams, dihydromaltophilin (**89**) exhibited the strongest activity, with a MIC value of 2.56 μM and the replacement of 14-OH with a carbonyl group resulted in decreased activity of xanthobaccin A (**92**), suggesting the 14-OH is critical for anti-BCG activity. On the other hand, the absence of 25-OH also led to the decrease of anti-BCG activity in compounds **90** and **91**. The loss of anti-BCG in compound **93** up to concentration of 80 μg/mL further confirmed the essential role of two hydroxyl groups on BCG growth inhibition. None of the isolated compounds gave positive results on the other tested models.

Table 7.6 Antimicrobial activities of compounds **89 – 95**

MIC (μM)	Organism (Strain)				
	BCG	SA	MRSA	BS	PA
89	2.56	NA	NA	NA	NA
90	9.92	NA	NA	NA	NA
91	10.2	NA	NA	NA	NA
92	20.4	NA	NA	NA	NA
93	NA	NA	NA	NA	NA
94	NA	NA	NA	NA	NA
95	NA	NA	NA	NA	NA
Control	0.37[a]	0.7[b]	0.7[b]	0.35[b]	3[c]

[a] isoniazid [b] vancomycin [c] ciprofloxacin * NA: not active

Figure 7.8 Comparison of NMR fingerprint between target fraction and compounds of MS110167.

Many polycyclic tetramate macrolactam (PTM) type compounds also displayed anti-

fungal activity[195-199, 202-204], especially for compound **89**, which has been widely studied as a biocontrol agent for fungal infestations of agricultural crops.[205] This is the first report of anti-microbial activity of PTM compounds, providing a new path for the practical application of PTMs in drug discovery.

Comparison of NMR spectra between pure compounds and target fraction confirmed all the target signals were identified (Figure 7.8).

7.4 Large-scale isolation of anti-TB compounds from LS120194_Medium3 _14d _Supernatant_fr.3

Desert actinomycete LS120194 was isolated from a desert sample collected from Taklimakan Desert and was identified as a *Streptomyces* sp. based on 16s rRNA phylogenetic analysis (Figure S7).

During our exploration for bioactive compounds from unique environments-derived actinomycetes, the desert strain LS120194 was found to be active against BCG (chapter 3). Guided by the OSMAC approach, 320 fractions were further generated from this strain under 32 different culture conditions. Anti-BCG screening resulted in 95 fractions (29.09%) showing activity against BCG with MIC values less than 50 μg/mL in a cell-based screening assay (Chapter 6). The third fraction of supernatant extract generated from LS120194 cultured in the medium 3 for 14 days (LS120194_Medium3_14d _Supernatant_fr.3) is one of the most active fractions with MIC values of 1.56 μg/mL (Figure S8).

A comparison between 320 fractions was conducted with the use of PCA. Multiple fractions derived from different conditions were highlighted as outliers, some of which were highlighted as outliers in more than one chemical region, such as the most active fraction LS120194_Medium3_14d_Supernatant_fr.3 (Figure S9). Visual inspection of the corresponding ^{1}H NMR spectra clearly displayed the metabolic profile changed significantly in the fraction of LS120194_Medium3_14d_Supernatant _fr.3 compare to those obtained under the initial condition (Figure 7.9).

With the aim of identifying metabolites produced under the specific condition indicated by PCA, a large scaled fermentation (4 L) of LS120194 cultured in medium 3 for 14 days was carried out. Further chemical investigation of the target fraction 3 afforded 9 compounds (**96 – 104**).
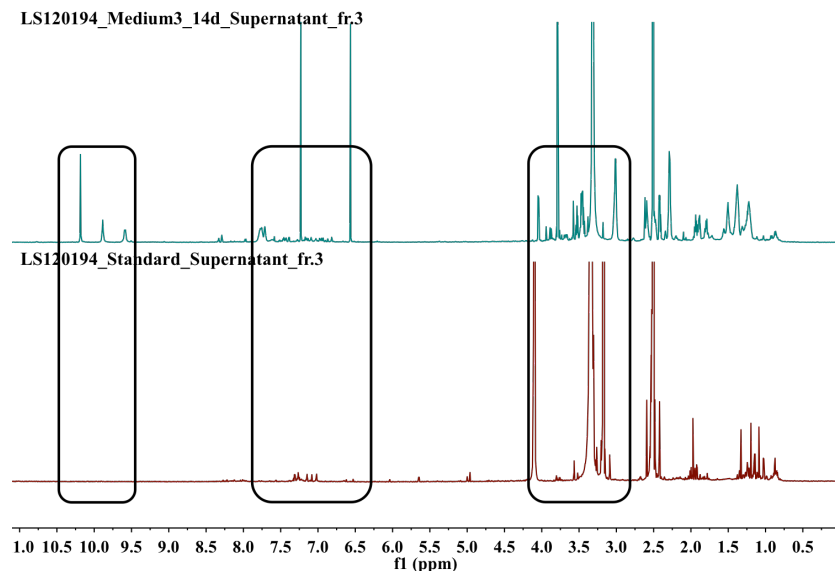
Figure 7.9 ^{1}H NMR spectra of the outlier LS120194_Medium3_14d_Supernatant_fr.3 indicated by PCA result of regions 2, 3 and 4 and the corresponding standard fraction.

Compound **96** was obtained as a white amorphous powder. The ^{1}H NMR spectrum of the **96** showed signals for one methyl group at δ_H 1.73 (3H, d, J=1.18 Hz), three olefinic protons at δ_H 7.85 (1H, d, J=8.24 Hz), 5.64 (1H, dd, J=2.25, 8.05 Hz) and 7.25 (1H, dq, J=5.77, 1.16 Hz), two oxygenated methine groups at δ_H 4.23 (1H, m) and 3.78 (1H, m), three down-field exchangeable singlets at δ_H 10.58, 11.00 and 11.29, and one signal at δ_H 6.16 (1H, dd, J=6.08, 7.55) corresponding to a secondary carbon attached to two oxygen or one oxygen and one nitrogen. In addition, the analysis of the ^{13}C NMR spectrum on the basis of the HSQC and HMBC correlations clearly showed the occurrence of one fully substituted olefinic carbon (δ_C 108.2), three olefinic methine (δ_C 141.0, 102.4 and 138.3), three secondary oxygenated carbons (δ_C 71.1, 84.7 and 88.2), one primary oxygenated carbon (δ_C 62.0), and four down-field fully substituted carbons at δ_C 151.8, 151.8, 163.4 and 165.3 ppm. Two fragments were established from the analysis of the COSY and HMBC correlations (Figure 7.10). The structure of sugar moiety in fragment A was deduced from the analysis of the COSY correlations between protons at δ_H 6.16, 2.10, 4.23, 3.78 and 3.58 ppm (H-6 – H-10). The long-range correlations from the proton signal at δ_H 7.85 (H-3) to the carbon resonances at δ_C 151.1 (C-2), 163.4 (C-5) and 84.7 (C-6) and from NH proton at δ_H 11.29 (H-1) to δ_C 102.4 (C-4) and 163.4 (C-5) revealed the connectivity of the pyrimidinedione and sugar rings of fragment A. Similarly, the HMBC spectrum indicated correlations between the proton signals at δ_H 7.25 (H-15) and

the carbon resonances at δ_C 151.8 (C-11), 165.3 (C-13), 108.2 (C-14), and 12.0 (CH$_3$-17), from NH proton at δ_H 11.00 (N-12) to carbons at δ_C 165.3 (C-13) and 108.2 (C-14), and from another NH proton at δ_H 10.58 (N-16) to 108.2 (C-14) establishing the structure of fragment B. No additional HMBC correlations were observed either from the two NH protons to carbons (C-8 or C-10) of fragment A or from H-8 or H-10 to carbons of fracgment B. In particular, the carbon chemical shift of C-11 occurs at 151.8 ppm. All 14 carbons displayed in ^{13}C spectrum are accounted for in the two fragments. Fragments could be joined between the diol (fragment A) and a carbonyl (C-11 in fragment B).

Fragment A **Fragment B**
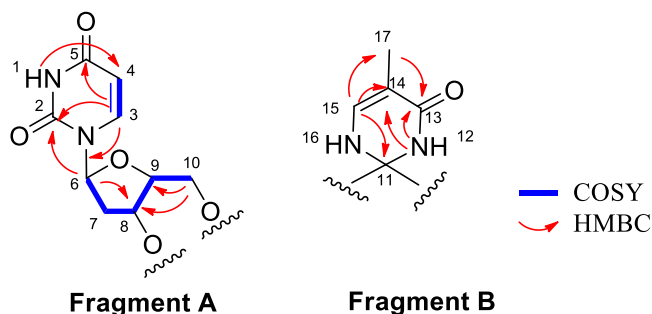
Figure 7.10 Structures and observed COSY and key HMBC correlations of two fragments from **96**.

Table 7.7 ^{1}H and ^{13}C (800 MHz) NMR data of **96** in DMSO-d_6

Position	δ_C, mult	δ_H, mult (J in Hz)	COSY	HMBC
1		11.29 s		4, 5, 6
2	151.1 C			
3	141.0 CH	7.85 d (8.24)	4	2, 4, 5, 6
4	102.4 CH	5.64 dd (2.25, 8.05)	3	3, 5
5	163.4 C			
6	84.7 CH	6.16 dd (6.08, 7.55)	7, 7'	2, 3, 8
7	40.1 CH$_2$	2.10 m	6, 7', 8	6, 8, 9
		2.07 m	6, 7, 8	6, 8, 9
8	71.1 CH	4.23 m	7, 7', 9	6, 10
9	88.2 CH	3.78 m	8, 10, 10'	6, 7, 8
10	62.0 CH$_2$	3.58 dd (3.86, 11.83)	9, 10'	8, 9
		3.54 dd (3.91, 11.88)	9, 10	8, 9
11	151.8 C			
12		11.00 s		13, 14
13	165.3 C			
14	108.2 C			
15	138.3 CH	7.25 dq (1.16, 5.77)	16, 17	11, 13, 14, 17
16		10.58 s	15	14
17	12.0 CH$_3$	1.73 d (1.18)	15	13, 14, 15

This conclusion was supported by COSY and HMBC correlations. All this evidence allowed to deduce that **96** was characterized by the generation of another six-membered ring between fragment A and B, and the occurrence of ketal functions on C-11. The structure of **96** was further confirmed by HRESIMS analysis. The HRESIMS spectrum of **96** acquired in positive mode gave the highest ion peak at m/z 251.0644 ([M+Na]$^+$, calculated for 251.0638) and was in accordance with the molecular formula $C_9H_{12}N_2O_5$, corresponding to fragment A with two hydroxyl groups at C-9 and C-10 (Figure 7.11).
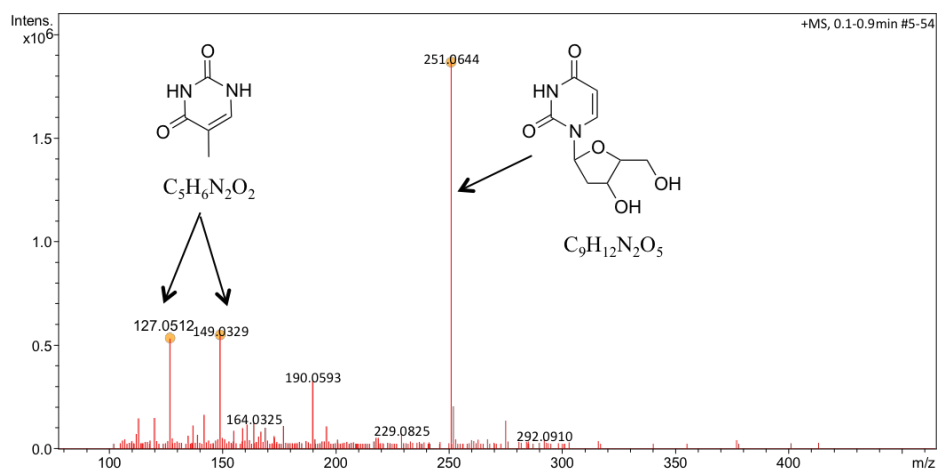
Figure 7.11 HRESIMS spectrum obtained from **96**.

The HRESIMS analysis also gave two intense peaks at m/z 127.0512 ([M+H]$^+$, calculated for 127.0502) and 149.0329 ([M+Na]$^+$, calculated for 149.0321), consistence with molecular formula of $C_5H_8N_2O$, which is corresponding to fragment B with a carbonyl at C-11. Therefore, the structure of **96** was deduced as shown in Figure 7.12. However, correlations between NH-12, NH-16 and H-8, H-10 were not observed in 1D and 2D ROESY spectra, the stereochemistry of C-11 was not determined.
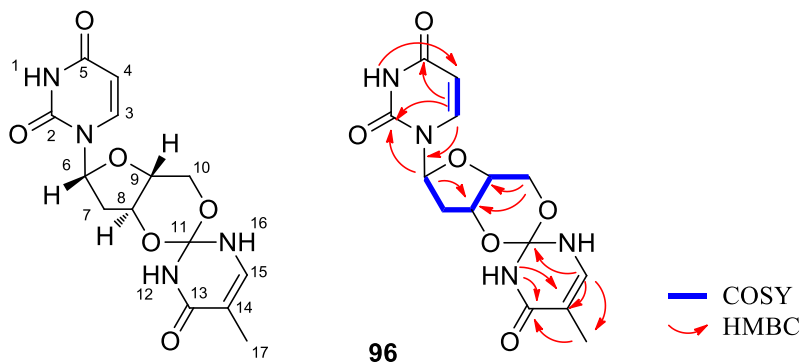
Figure 7.12 Structures and observed COSY and key HMBC correlations of **96**.

97 was isolated as white powder. Its molecular formula was established as $C_{10}H_{12}N_4O_5$ by HRESIMS (m/z 269.0886 $[M+H]^+$, calculated for 269.0880). Interpretation of the ^1H NMR data of **97** revealed the presence of two olefinic protons at δ_H 8.33 (s) and 8.07 (s) ppm, six oxygenated methine protons at δ_H 5.87 (d, J=5.8 Hz), 4.49 (t, J=3.56 Hz), 4.13 (t, J=3.86 Hz) and 3.94 (dd, J=3.86, 7.44 Hz), 3.66 (dt, J=3.49, 11.86 Hz) and 3.56 ppm (dt, J=4.66, 12.07 Hz) and three hydroxyl groups at δ_H 5.51, 5.23 and 5.12 ppm, suggesting the existence of a sugar moiety (Table 7.8). The ^{13}C NMR spectrum displayed 10 carbons, including 4 olefinic carbons at δ_C 139.2, 146.5, 148.7 and 124.9 ppm, one up-field carbonyl at δ_C 157.2 ppm and five oxygenated carbons belonging to the sugar part at δ_C 87.9, 74.5, 70.7, 86.1 and 61.8 ppm. Extensive analysis of 2D NMR spectra suggested **97** consisted of a purine base and one ribose moiety. The assignment of the purine base was confirmed by the HMBC correlations from H-2 (8.07 ppm) to C-3 (148.7 ppm) and C-5 (157.2 ppm), along with correlations from H-6 (8.33 ppm) to C-3 (148.7 ppm) and C-4 (124.9 ppm). The HMBC spectrum showed correlations from H-1' (5.87 ppm) to C-3 (148.7 ppm) and C-6 (139.2 ppm), indicating that the sugar component is connected to the middle N between C-3 and C-6. Thus, the structure of compound **97** was dereplicated as shown in Figure 7.13.

97

98: R_1=H, R_2=OCH$_3$
99: R_1=CH$_3$, R_2=H
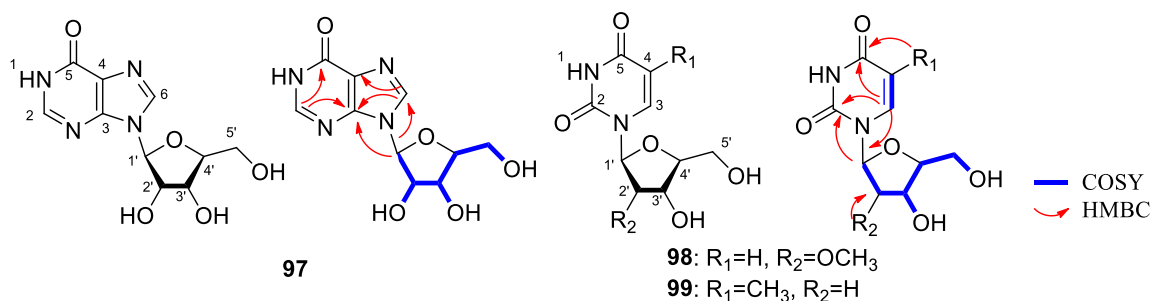
— COSY
↩ HMBC

Figure 7.13 Structures and observed COSY and key HMBC correlations of **97 – 99**.

98 was obtained as white powder. Its molecular formula was determined as $C_{10}H_{14}N_2O_6$ by HRESIMS (m/z 259.0835 $[M+H]^+$, calculated for 259.0825). 1D and 2D NMR data of **98** was highly similar with that of **97**, except for the less of two olefinic carbons and one additional carbonyl at δ_C 151.6 ppm and the appearance of one additional methoxyl group in **98** (Table 7.8). The position of the carbonyl and methoxyl groups was confirmed by HMBC correlations from H-3 (8.07 ppm) to C-2 (121.0 ppm) and H-6' (3.36 ppm) to C-2' (82.7 ppm). Therefore, structure of **98** was determined as shown in Figure 7.13.

Table 7.8 ^{1}H and ^{13}C (800 MHz) NMR data of **97** and **99** in DMSO-d_6

Position	97 δ$_C$, mult	97 δ$_H$, mult (J in Hz)	98 δ$_C$, mult	98 δ$_H$, mult (J in Hz)	99 δ$_C$, mult	99 δ$_H$, mult (J in Hz)	COSY	HMBC
1		11.37 s		11.35 s		11.27 s		
2	146.5 CH	8.07 s	151.0 C		150.9 C			3, 5
3	148.7 C		140.4 CH	7.94 d (5.17)	136.6 CH	7.70 d (1.29)	4	2, 4, 5, 6, 1'
4	124.9 C		101.9 CH	5.66 d (7.98)	109.8 C		3	3, 5
5	157.2 C		163.6 C		164.2 C			
6	139. 2 CH	8.33 s			12.8 CH$_3$	1.78 d (1.13)		3, 4, 5
1'	87.9 CH	5.87 d (5.8)	86.0 CH	5.86 d (5.17)	84.2 CH	6.17 dd (5.98, 7.61)	2'	3, 6, 2'
2'	74.5 CH	4.49 t (5.52)	82.7 CH	3.79 t (5.06)	40.1 CH$_2$	2.09 m	1', 3', 2'-OH	1', 4', 6'
						2.06 m		1', 4', 6'
3'	70.7 CH	4.13 t (3.86)	68.3 CH	4.12 dd (4.96, 10.47)	70.9 CH	4.24 m	2', 4', 3'-OH	1'
4'	86.1 CH	3.94 dd (3.86, 7.44)	85.1 CH	3.86 dd (3.22, 6.75)	87.7 CH	3.76 dd (3.79, 6.84)	3', 5', 5"	3'
5'	61.8 CH$_2$	3.66 dt (3.49. 11.86)	60.5 CH$_2$	3.64 m	61.8 CH$_2$	3.60 m	4', 5", 5'-OH	3', 4'
		3.56 dt (4.66, 12.07)		3.57 dd		3.55 m	4', 5', 5'-OH	3', 4'
6'			57.6 CH$_3$	3.36 s				2'
2'-OH		5.51 s					2'	
3'-OH		5.23 s		5.16 s		5.23 d (4.31)	3'	2', 3', 4'
5'-OH		5.12 s		5.14 d		5.02 t (5.19)	5'	4', 5'

Compound **99** was obtained as a white solid. The molecular formula was determined to be $C_{10}H_{14}N_2O_5$ from the HRESIMS (m/z 265.0792 [M+Na]$^+$, calculated for 265.0795). ^{1}H and ^{13}C NMR data obtained for **99** showed high similarity to the data obtained for compound **98** (Table 7.8). The major difference in the NMR data for **98** and **99** was that the H-4 proton at δ$_H$ 5.66 ppm in **98** was replaced by a methyl group (δ$_H$ 1.78, d, J=1.13 Hz) and the absence of the methoxy group attached to C2' in **99**. Key HMBC correlation from methyl protons to C-3 (δ$_C$ 136.6), C-4 (δ$_C$ 109.8) and C-5 (δ$_C$ 157.2) confirmed the structure of **99** (Figure 7.13).

Compound **100** was obtained as a white solid. The positive ion [M+H]$^+$ at m/z 281.1134 (calculated for 281.1132) in the HRESIMS suggested the molecular formula of **100** as $C_{13}H_{16}N_2O_5$. ^{1}H NMR revealed 4 aromatic protons at δ$_H$ 7.18, 7.61, 8.00 and 8.56 ppm, 2 methine at δ$_H$ 4.39 and 5.18 ppm, two methyl groups at δ$_H$ 1.36 and 1.45 ppm and 2 exchangable protons with the same chemical shifts at δ$_H$ 11.94 ppm (Table 7.9). The ^{13}C NMR spectra displayed 13 resonances including three carbonyl groups at δ$_C$ 170.8, 174.8

and 169.5 ppm, six aromatic carbons (four methines and two fully substituted) at δ_C 117.0 ~ 140.6 ppm, two methines at δ_C 66.2 and 70.7 ppm, and two methyls at δ_C 20.6 and 18.6 ppm. The combined use of COSY and HSQC experiments clearly suggested the presence of a dipeptide composed by two alanine residues linked to the 1,6-symmetrical disubstituted phenyl ring moiety. The downfield shifted of the first alanine α-proton (H-9) at δ_H 4.39 ppm suggested that the aromatic moiety was attached to the free amino group of the alanine. This was confirmed by the long range C–H correlation observed in the HMBC spectrum between the carbonyl carbon C-7 at δ_C 170.8 ppm of the aromatic moiety and the H-9 proton of the arginine residue. Further long range C–H correlations observed in the HMBC from H-10 (1.36, d, J= 6.5 Hz) to C-9 (66.2) and C-11 (174.1), from H-14 (1.45, d, J= 6.5 Hz) to C-13 (70.7) and C-15 (169.5) allowed the identification of the three fragments (Figure 7.14).

Table 7.9 ^{1}H and ^{13}C (800 MHz) NMR data of **100 – 102** in DMSO-d_6

Position	100		101		102		COSY	HMBC
	δ_C, mult	δ_H, mult (*J* in Hz)	δ_C, mult	δ_H, mult (*J* in Hz)	δ_C, mult	δ_H, mult (*J* in Hz)		
1	131.5 CH	8.00 dd (1.1, 8.0)	131.7 CH	8.01 d (7.8)	126.2 CH	8.09 dd (1.5, 8.0)	2	3, 5, 7
2	123.4 CH	7.18 td (1.2, 8.4)	122.8 CH	7.12 t (7.5)	126.3 CH	7.49 td (0.9, 7.5)	1, 3	1, 4
3	134.6 CH	7.61 td (1.7, 8.0)	133.6 CH	7.52 t (7.7)	134.7 CH	7.80 td (1.5, 7.6)	2, 4	1, 4, 5
4	120.0 CH	8.56 dd (1.1, 8.5)	119.7 CH	8.67 d (8.6)	127.0 CH	7.59 d (8.0)	3	2, 5, 6
5	140.6 C		140.8 C		149.4 C			
6	117.0 C		118.7 C		121.1 C			
7	170.8 C		169.7 C		162.2 C			
8		11.94 s		12.32 s		12.20 s		
9	66.2 CH	4.39 q (6.8)	68.4 CH	4.14 q (6.7)	154.7 C		10	7, 10, 11
10	20.6 CH$_3$	1.36 d (6.5)	21.4 CH$_3$	1.33 d (7.2)	21.9 CH$_3$	2.39 s	9	9, 12
11	174.1 C		174.8 C					
12		11.94 s						
13	70.7 CH	5.18 q (7.1)					14	11, 14, 15
14	18.0 CH$_3$	1.45 d (6.6)					13	13, 15
15	169.5 C							
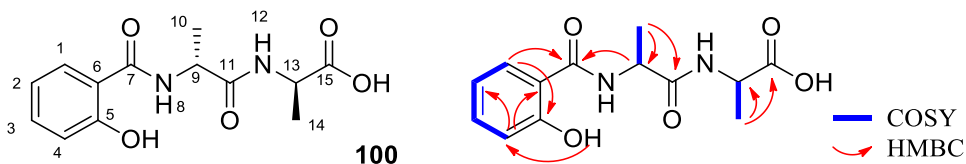
Figure 7.14 Structures and observed COSY and key HMBC correlations of **100**.

Compound **100** was a synthetic peptide for anti-inflammatory research. It improved the activity and selectivity of nonsteroidal anti-inflammatory drugs (NSAID) by the formation of hydrogels via conjugation.[206] This is the first report of this compound as a natural product.

Compound **101** was obtained as a white solid. The HRESIMS showed a $[M + Na]^+$ peak at m/z 232.0579 (calculated for 232.0580), consistent with the molecular formula of $C_{10}H_{11}NO_4$. The 1D and 2D NMR data obtained for **101** (Table 7.9) was nearly identical to the data obtained for **100**. The main difference between them is the absence of one NH, one methine (C-13), one methyl (C-14) and one carbonyl (C-15) group in **101**. The key HMBC correlations from H-1 (8.01, d, J=7.8 Hz) to C-5 (140.8 ppm) and C-7 (169.7 ppm), H-9 (4.14, q, J=6.7 Hz) to C-7 (169.7 ppm) and C-11 (174.8 ppm) confirmed the structure of **100** as shown in Figure 7.24.
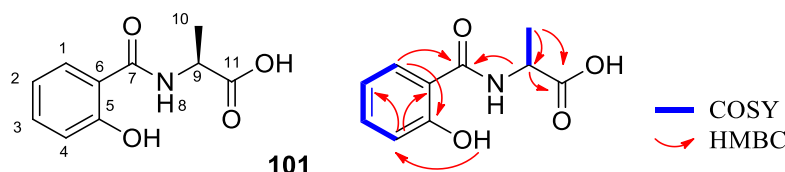
Figure 7.15 Structures and observed COSY and key HMBC correlations of **101**.

102 was isolated as white powder. Its molecular formula was determined as $C_9H_8N_2O$ based on HRESIMS result (m/z 161.0715 $[M+H]^+$, calculated for 161.0709). Similar with **100** and **101**, the presence of one 1, 2-di-substituted benzene moiety in **102** was evident from analysis of the 1D and 2D NMR data (Table 7.9). The assignments of the pyrimidine moiety were further confirmed by the HMBC correlations from H-1 (8.09, dd, J= 1.5, 8.0 Hz) to C-5 (149.4 ppm) and C-7 (162.2 ppm) and from H-10 (2.39 s) to C-9 (154.7 ppm). Thus, the structure of **102** was dereplicated as shown in Figure 7.16.
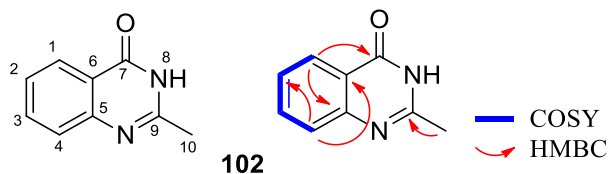
Figure 7.16 Structures and observed COSY and key HMBC correlations of **102**.

Compound **103** was obtained as yellow powder. The molecular formula was determined as $C_{13}H_{14}N_2O_4$ by HRESIMS (m/z 263.1028, $[M+H]^+$, calculated for 263.1026). Thirteen carbon signals were observed in the ^{13}C NMR spectrum and identified by HSQC (Table

7.10) as two aromatic methine carbons at δ_C 112.8 and 108.0 ppm, three sp3 methylenes at δ_C 26.1, 23.6 and 47.1 ppm, one methine at δ_C 56.8 ppm, one methoxy at δ_C 36.2 ppm, and six fully substituted carbons, including two carbonyls at δ_C 164.8 and 170.9 ppm, presumed to belong to amides based on their chemical shifts. The ^{1}H NMR spectrum showed an exchangeable proton signal at δ_H 10.20 ppm and two aromatic proton singlets at δ_H 6.56 and 7.23 ppm, indicating the presence of a 1,2,4,5-tetrasubstituted aromatic ring. In the aliphatic region, a methoxy singlet at δ_H 3.78 ppm and a methine doublet of doublets at δ_H 4.04 were observed. The latter signal showed couplings in the COSY spectrum with two multiplets at δ_H 2.48 and 1.94 assigned to the methylene protons at C-1. Further couplings were seen between the C-2 methylene protons at δ_H 1.88 and 1.79 and the C-1 methylene proton signals and C-3 methylene protons at δ_H 3.53 and 3.44 ppm. The HMBC experiments revealed correlations from the methoxy protons and the two aromatic protons (H-6 and H-9) to C-7 (δ_C 145.0). The H-6 proton exhibited further couplings with C-9a (δ_C 130.9), C-5a (δ_C 117.8), and C-5 (δ_C 164.8), and H-9 showed additional correlations with C-5a. Together, these data and the chemical shifts of C-9 and C-9a allowed the placement of one nitrogen atom at C-9a. Key correlations were seen between H-11a and C-11 (δ_C 170.9), C-2 (δ_C 23.6), and C-5. These observations and the chemical shift of H-11a support the placement of the second nitrogen atom between C-11a and the C-5 carbonyl and complete the core structure to be pyrrolobenzodiazepine. The protons on C-1 also correlated with C-11 and C-2. On the basis of the above data, the structure was constructed for **103** as shown in Figure 7.17.

103: R=H
104: R=OH
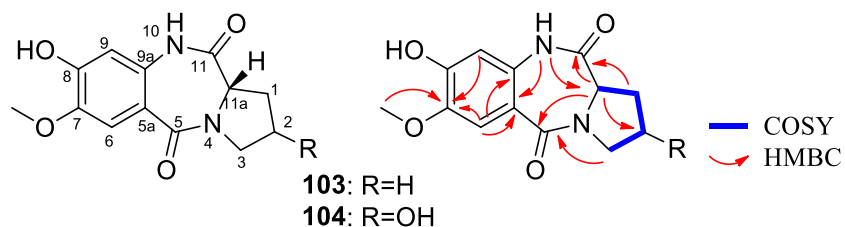
Figure 7.17 Structures and observed COSY and key HMBC correlations of **103** and **104**.

Compound **104** was obtained as yellow powder. Its molecular formula was determined to be $C_{13}H_{14}N_2O_5$ according to HRESIMS result (m/z [M+H]$^+$ 279.0979, calculated for 279.0975). A close similarity in ^{1}H and ^{13}C NMR spectra of **104** to **103** suggested that the skeletons of the compounds are similar (Table 7.10). The methylene C-2 of **103** was replaced by a oxygenated methine carbon at δ_C 68.0 ppm (δ_H 4.27) and one additional

hydroxyl signal was observed in **104** at δ_H 4.16 ppm. The hydroxyl group was assigned by the COSY correlation with H-2 and HMBC correlations to C-1 (34.3) and C-3 (54.9). Thus, the structure of **104** was identified as shown in Figure 7.17.

Table 7.10 ^{1}H and ^{13}C (800 MHz) NMR data of **103** and **104** in DMSO-d_6

Position	103		104		COSY	HMBC
	δ_C, mult	δ_H, mult (J in Hz)	δ_C, mult	δ_H, mult (J in Hz)		
1	26.1 CH$_2$	2.48 m	34.3 CH$_2$	2.53 m	1', 2, 2', 11a	2, 3, 11
		1.94 m		2.16 m	1, 2, 2', 11a	2, 11, 11a
2	23.6 CH$_2$	1.88 m	68.0 CH	4.27 m	1, 1', 2', 3	2, 11a
		1.79 m			1, 1', 2, 3	1, 3
3	47.1 CH$_2$	3.53 m	54.9 CH$_2$	3.76 dd (5.25, 11.90)	2, 2'	1, 11a
		3.44 m		3.22 dd (3.66, 11.96)	2, 2'	2
5	164.8 C		165.7 C			
5a	117.8 C		117.1 C			
6	112.8 CH	7.23 s	113.0 CH	7.23 s		5, 5a, 7, 8, 9a
7	145.0 C		144.9 C			
8	150.3 C		150.7 C			
9	108.0 CH	6.56 s	108.2 CH	6.58 s		5, 5a, 7, 8, 9a
9a	130.9 C		131.4 C			
10		10.20 s		10.22 s		5a, 9, 9a, 11, 11a
11	170.9 C		171.1 C			
11a	56.8 CH	4.04 dd (1.96, 7.85)	56.3 CH	4.08 dd (4.16, 8.80)	1, 1'	1, 2 ,3, 5, 11
7-OCH$_3$	56.2 CH$_3$	3.78 s	56.1 CH$_3$	3.78 s		7
2-OH				4.16 m	2	1, 2, 3

Compound **103** and **104** were synthesized as potent anti-tumor compounds in cancer drug discovery.[207, 208] This is the first report of these two pyrrolobenzodiazepine (PBD) compounds as natural products. The mechanism of action of this type of compounds was reported to form a covalent aminal bond with N2-amino group of guanine via interaction with the minor groove of DNA.[209] The unique mechanism is particular attractive for TB drug discovery as only four approved TB drugs exhibited DNA related mechanisms. It may provide an option to overcome the burden caused by MDR-TB.

The NMR fingerprint of the target fraction (LS120194_Medium3_14d_Supernatant_fr.3) and pure compounds (**96 - 104**) were compared in Figure 7.18. All the target signals were identified in pure compounds.
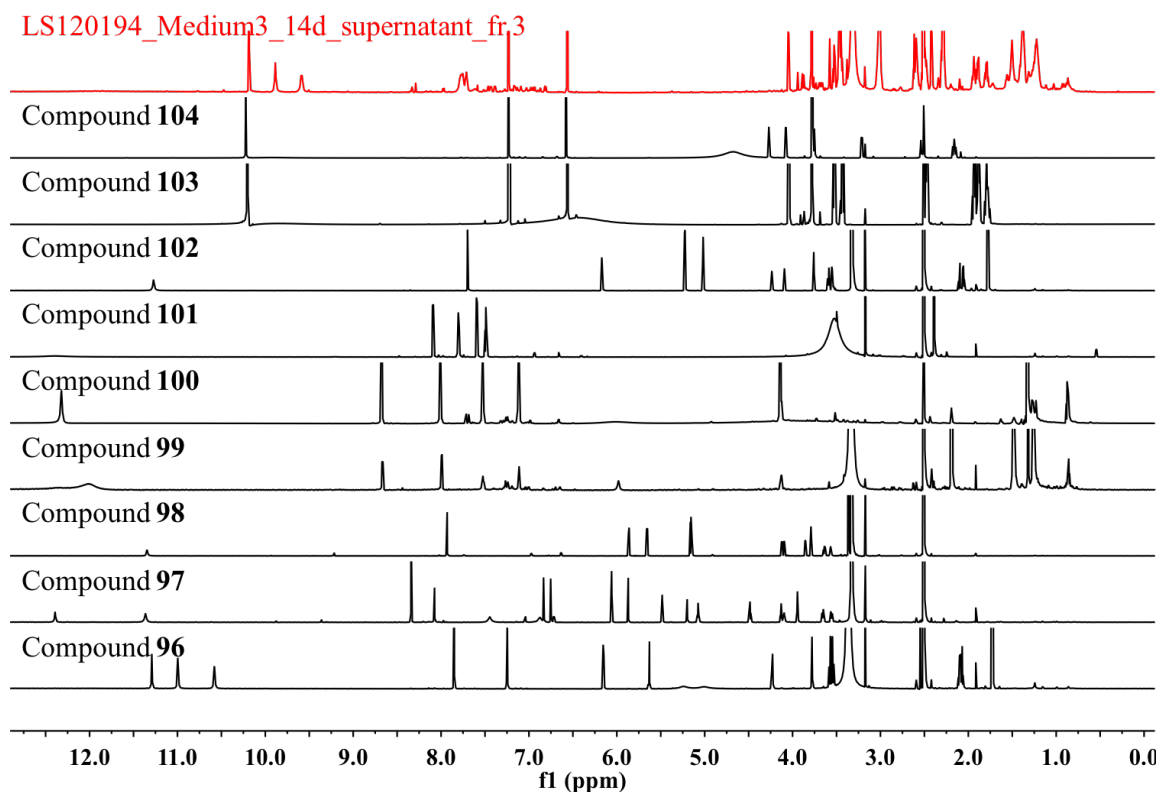
Figure 7.18 Comparison of NMR fingerprint between target fraction and compounds of LS120194.

7.5 Conclusion

Strains MS110104, MS110167 and LS120194 were investigated as part of the OSMAC project, which was designed to trigger the expression of silent genes and thus increase the numbers of secondary metabolites from one single strain through the variation of multiple culture conditions (chapter 6). The three strains were found to exhibit strong anti-TB activities and unique NMR signals under specific culture conditions. Large-scale fermentation and chemical investigation of the 3 target fractions resulted in the identification of 20 compounds, including 1 novel compound and 3 synthetic compounds that were isolated as natural products for the fist time.

Among the 11 compounds that have been evaluated for anti-bacteria activities, 8 of them showed anti-BCG activity with MIC values from 5 to 40 μg/mL (Table 7.11). The distinct structure–activity correlations were elucidated for compounds **89** – **93** based on these results. Both 14-OH and 25-OH groups were beneficial to enhance anti-BCG activity of the PTM compounds. The absence of the two groups decreased the activity to various degrees.

Table 7.11 Antimicrobial activities of compounds **85 – 95**

MIC (*μ*M)	Organism (Strain)				
	BCG	SA	MRSA	BS	PA
85	12.44	62.2	NA	NA	NA
86	12.72	63.6	NA	NA	NA
87	13	65	NA	NA	NA
88	24.16	60.4	NA	NA	NA
89	2.56	NA	NA	NA	NA
90	9.92	NA	NA	NA	NA
91	10.2	NA	NA	NA	NA
92	20.4	NA	NA	NA	NA
93	NA	NA	NA	NA	NA
94	NA	NA	NA	NA	NA
95	NA	NA	NA	NA	NA
Control	0.37[a]	0.7[b]	0.7[b]	0.35[b]	3[c]

This study has also led to the identification of two types of compounds that displayed unique mechanisms of action. Bafilomycins (**85 – 88**) were reported as potent inhibitors of ATPase, which is essential to mycobacteria and most importantly, apart from one drug candidate TMC207, there's no current anti-TB with the similar target. The second type of compound is pyrrolobenzodiazepines (**103** and **104**), which have been reported to target DNA. These compounds may represent useful scaffolds for future development as antituberculars to treat drug-sensitive as well as drug-resistant TB.

Chapter VIII Conclusion

Tuberculosis (TB) has several barriers to drugs, such as delivery to the lung, penetration of the granuloma, and permeation of the unusual cell wall of *M. tuberculosis,* which mainly consists of mycolic acid. Compounds must possess certain physicochemical properties to allow cell penetration and thus arrive at the target. Herein, we analyzed the four Lipinski's properties MW, log P, HBA and HBD as well as tPSA and numbers of rotation bonds of 50 secondary metabolites isolated from 4 endophytes, 2 desert strains and 5 marine strains (Table 8.1).

Table 8.1 Physicochemical profiling of the 50 isolated natural products

Compounds	Mol Weight	LogP	H bond acceptors	H bond donors	TPSA	Rotatable bonds	Number of violations	Anti-BCG (μM)	Source
55	388.42	5.54	5	1	80.67	3	1	258	Endophytes
56	390.44	5.9	5	1	80.67	4	1	NA	Endophytes
57	366.41	6.91	5	2	91.67	5	1	17	Endophytes
58	364.4	6.55	5	2	91.67	4	1	62	Endophytes
59	404.42	4.38	6	2	100.9	3	0	NA	Endophytes
60	404.42	4.35	6	2	100.9	3	0	NA	Endophytes
61	526.54	1.52	9	5	155.14	4	1	NA	Marine strains
62	508.52	2.6	8	4	134.91	4	1	6	Marine strains
63	510.54	2.82	8	4	134.91	4	1	25	Marine strains
64	489.65	4.66	6	3	127.85	2	0	1.5	Desert strains
65	945.11	-0.39	19	12	315.21	9	3	106	Desert strains
66	295.3	-1.15	8	3	116.76	3	0	NA	Endophytes
67	283.33	0.95	2	2	65.2	2	0	14.15	Endophytes
68	279.25	2.39	4	2	67.79	0	0	27.9	Endophytes
69	307.26	2.9	5	2	84.86	0	0	30.7	Endophytes
70	210.27	0.93	3	2	57.53	0	0	NA	Endophytes
71	161.16	1.65	2	2	53.09	1	0	NA	Endophytes
72	176.18	2.6	4	2	66.24	0	0	NA	Endophytes
73	1,025.28	6.4	16	8	269.82	14	4	1.5	Marine strains
74	1,011.25	5.95	16	8	269.82	13	4	49	Marine strains
75	275.3	1.99	4	2	83.47	3	0	NA	Marine strains
76	291.3	1.2	5	3	103.7	3	0	343	Marine strains
77	281.35	0.9	4	2	83.47	3	0	NA	Marine strains
78	562.52	0.63	13	7	204.83	5	3	NA	Endophytes
79	416.38	1.86	9	5	145.91	3	0	NA	Endophytes

80	298.29	3.37	5	1	64.99	3	0	NA	Endophytes
81	145.16	1.78	1	1	32.86	1	0	NA	Endophytes
82	254.24	2.73	4	2	66.76	1	0	NA	Endophytes
83	270.24	3.08	5	3	86.99	1	0	NA	Endophytes
84	388.46	1.31	6	4	142.03	11	0	8	Endophytes
85	622.84	5.08	8	4	134.91	7	2	12.44	Marine strains
86	636.87	5.72	8	3	123.91	8	2	12.72	Marine strains
87	650.89	6.37	8	2	112.91	9	2	13	Marine strains
88	604.83	5.71	7	3	122.52	10	2	24.16	Marine strains
89	512.65	1.42	6	5	135.96	1	1	2.56	Marine strains
90	496.65	2.56	5	4	115.73	1	0	9.92	Marine strains
91	510.68	2.86	5	4	115.73	1	1	10.2	Marine strains
92	510.63	1.8	6	4	132.8	1	1	20.4	Marine strains
93	494.63	2.95	5	3	112.57	1	0	NA	Marine strains
94	943.13	0.95	18	11	294.98	9	3	NA	Marine strains
95	957.16	1.09	17	10	283.98	10	3	NA	Marine strains
96	336.3	0.14	7	3	118.23	1	0	ND	Desert strains
97	268.23	-2.48	7	4	129.2	2	0	ND	Desert strains
98	258.23	-1.77	6	3	108.33	3	0	ND	Desert strains
99	242.23	-1.12	5	3	99.1	2	0	ND	Desert strains
100	280.28	0.9	5	4	115.73	5	0	ND	Desert strains
101	209.2	1.44	4	3	86.63	3	0	ND	Desert strains
102	160.18	0.77	2	1	41.46	0	0	ND	Desert strains
103	262.27	1.02	4	2	78.87	1	0	ND	Desert strains
104	278.26	-0.13	5	3	99.1	1	0	ND	Desert strains
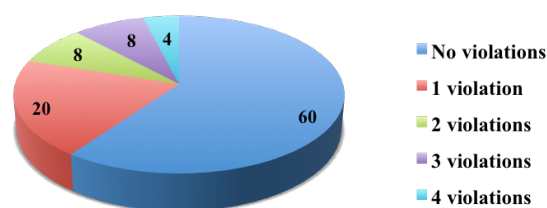
Figure 8.1 Pie chart presentation of the percentage of 50 isolated compounds obeying or violating Lipinski's rule of five.

The results indicated that 80% of the isolated natural products obeyed Lipinski's rule-of-five with less than one violation (Figure 8.1), in terms of MW ≤ 500 Da (68%), log P ≤ 5 (80%), HBA ≤ 10 (88%), HBD ≤ 5 (90%) number of rotation bonds ≤ 10 (90%) and tPSA ≤ 140 Å (84%). The existence of a sugar moiety in the molecules provided more *O-*

containing functionalities and hydroxyl groups, which might account for the violations of HBA and HBD values, as well as molecular weights. This is consistent with the result compared to other source, of secondary metabolites sourced from marine environments (chapter 2).

Anti-BCG activities of isolated compounds are displayed in Table 8.1. The six most active compounds **64** (1.5 μM), **73** (1.5 μM), **89** (2.56 μM) and **90** (6 μM), **62** (8 μM), and **84** (9.92 μM) with MIC values less than 10 μM were identified from a desert strain LS120167, a marine actinomycete MS110109, a marine actinomycete MS110167, a marine actinomycete MS110154 and an endophyte ES120127, suggesting the potential of producing anti-TB natural products from microbes derived from the three unique sources. Except for compound **73** violating all the six rules in Table 8.1, compounds **64**, **84** and **90** perfectly obeyed Ro5 as well as rules of tPSA and rotatable bonds, indicating their high potential for further investigation to be drug candidates. Compounds **89** and **62** only violated the molecular weight rule with 512 and 508 Da, respectively.
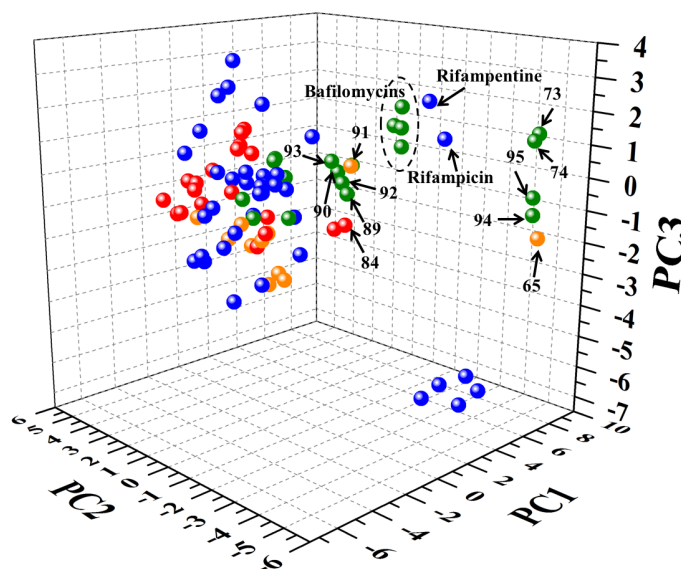
Figure 8.2 Distribution of 39 anti-TB drugs and 50 isolated natural products in ChemGPS-NP chemical space. 39 anti-TB drugs, 20 natural products isolated from endophytes, 11 natural products isolated from desert strains and 19 natural products isolated from marine strains were shown in blue, red, orange and green. PC1 representing molecular size, PC2 representing aromatic and conjugation related properties; PC3 representing lipophilicity, polarity, and H-bond capacity.

ChemGPS-NP was also used to compare the distribution of the 50 isolated compounds and 39 anti-TB drugs in physicochemical space (Figure 8.2). Around 70% of the 50

isolated natural products are positioned within similar drug-like chemical space as the anti-TB compounds with relative medium molecular size (PC1), high aromaticity (positive direction of PC2) and various polarity (PC3), which means that these compounds share similarities in the described parameters. Two types of large molecules, triterpenes **65**, **73**, and **74**, and macrodiolides **94** and **95** occupied at the extremely positive direction of PC1 and bafilomycins identified from marine strain MS110104 clustered with microbe derived TB drugs rifampicin and rifampetine. Particularly, compounds (**64**, **84**, **89**, **90**, **91** and **92**) showing strong anti-BCG activities with MIC values from 1.5 to 20 μM and complete compliance to Ro5 clustered together, indicating the potential relationship between anti-TB activity and physico-chemical properties of microbial natural products.

The 50 isolated compounds were also evaluated within the 27-region chemical space defined in chapter 2 (Figure 8.3). The regions containing most compounds were found to be regions 5 (7 compounds), 10 (8 compounds) and 23 (9 compounds). The most active compounds were located in regions 5 (**64**, **84**, **89** and **90**), 9 (**73**) and 23 (**62**). In particular, region 5 with 4 active compounds only holds one current anti-TB drug meropenem, which is a cell wall synthesis inhibitor.
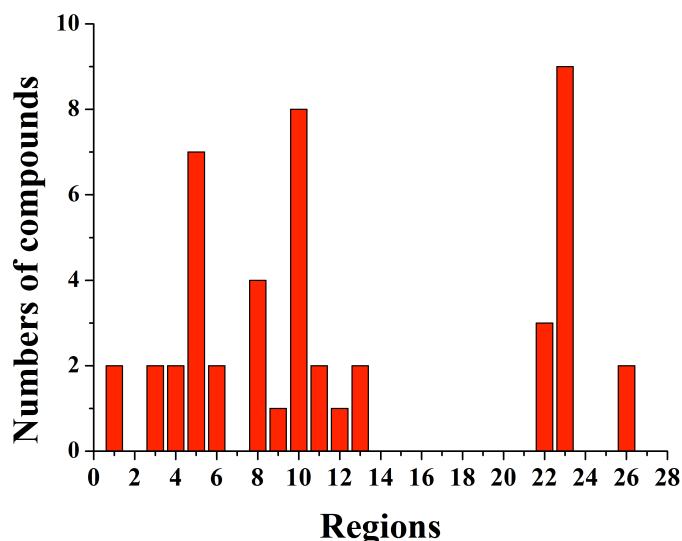
Figure 8.3 Histogram of the percentage of 50 isolated NPs in each region.

Filtering these compounds using ED calculation to identify near neighbors with 39 known TB dugs produced 18 (36%) prioritized structures with ED < 2 to at least one drug (Figure 8.4). The shortest distances were found between compounds **98** and **99** to anti-TB

drugs clavulanic acid, namely 0.87 and 0.98. New compound **96** showed short EDs with more than one drug neighbor (Figure 8.4), suggesting its high similarity with active drug in physico-chemical space. Besides **96**, anti-TB activities of compounds **97-104** have not been determined yet. Thus, the prediction model developed in chapter 2 was applied to these compounds. Except for compound **97**, all the other untested compounds showed a short distance to at least one anti-TB drug with EDs < 2, indicating the high chances for these compounds to be active against TB.

anti-TB drugs
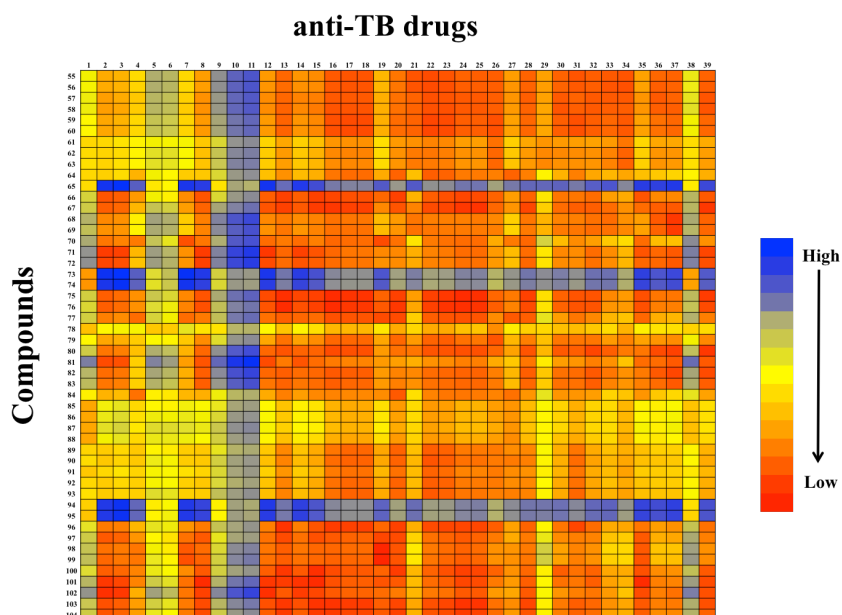
Figure 8.4 EDs between 50 isolated NPs and 39 TB drugs and candidates.
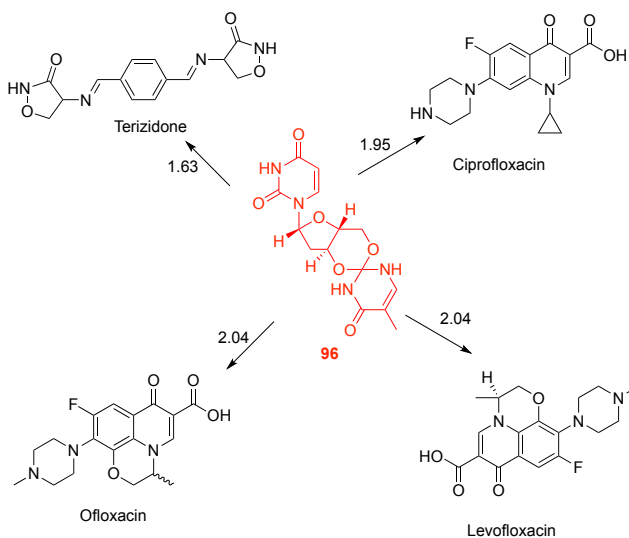
Figure 8.5 Structures of new compound **96** and its drug pairs. EDs are given in black numbers.

In conclusion, the investigation concentrated on microbial natural products that had significant activities on TB was investigated by 3 types of methods in this thesis, namely bioassay-guided isolation (chapter 4), NMR fingerprinting-guided isolation (chapter 5) and NMR fingerprinting conjugated with OSMAC strategy and PCA - guided isolation (chapter 6 and 7). The first method resulted in the identification of 6 compounds, including 4 new compounds from an endophyte Y3111. Three compounds exhibited moderate anti-BCG activity. Applying the second method on 21 active fractions, representing 8 strains, led to the identification of 24 natural products, including one new compound. Four compounds displayed strong anti-TB activity with MIC values of 1.5, 1.5, 6 and 8 μM, suggesting the reliability for obtaining active constituents under the direction of NMR fingerprinting. The low yield of many secondary metabolites in actinomycetes led to failure in identifying more active but minor compounds. Therefore, the third method was developed to induce or enhance the expression levels of secondary metabolite biosynthesis gene clusters by screening the metabolomic profiles of 13 strains under different culture media or stress conditions. Rigorous evaluation with PCA of the NMR fingerprint dataset consisting of 4160 fractions identified 37 outlier fractions with unique NMR information. Detailed investigation of 3 of the most active fractions resulted in the isolation of 20 compounds, including one new compound with a rare skeleton and 8 out of 11 evaluated compounds exhibited strong anti-BCG activities from 2.56 to 24.16 μM. The discovery of a new structure and active compounds proved that the combined NMR-OSMAC fingerprinting can be employed as a beneficial strategy for the selection of ideal candidates for further investigation on tuberculosis.

Experimental section

General Experimental Procedures. Optical rotations were recorded on a JASCO P-1020 polarimeter (10 cm cell). UV spectra were obtained on a Cary 50 spectrophotometer. NMR spectra were recorded in DMSO-d_6 (δ_H 2.50 and δ_C 39.5), CD$_3$OD (δ_H 3.35, 4.78 and δ_C 49.3), C$_6$D$_6$ (δ_H 7.26 and δ_C 128.0) or CDCl$_3$ (δ_H 7.24 and δ_C 77.0) at 30 °C on a Varian INOVA 600 MHz spectrometer equipped with a triple-resonance cold probe or at 25 °C on a Bruker Avance HDX 800 MHz spectrometer equipped with a TCI cryoprobe. The low-resolution LC-MS was recorded on a Waters LCMS system equipped with a Luna C$_{18}$ column (3 μm, 100 Å, 50 × 4.6 mm), a PDA detector, and a ZQ ESI mass spectrometer. High-resolution mass spectra (HRESIMS) were recorded on a Bruker Daltonics SolariX 12 T Fourier transform mass spectrometer. Resin HP-20 (Diaion, Japan) and ODS-A (YMC, Japan) were used for purification. The HPLC system included a Waters 600 pump fitted with a 996 photodiode array detector and Gilson FC204 fraction collector. A Thermo Electron Betasil C$_{18}$ column (5 μm, 21.2 × 150 mm) and a Phenomenex Luna C$_{18}$ column (5 μm, 10 × 250 mm) were used for semi-preparative HPLC. All solvents used for extraction, chromatography, [α]$_D$, UV, and MS were Lab-Scan HPLC grade, and the H$_2$O was Millipore Milli-Q PF filtered.

Bacterial material. The endophytic strain Y3111 was isolated from stems of the TCM *H. souliei* collected in Zheduo Mountain of Kangding, Southwest of China (N30°03′, E101°49′, 4006 m). Endophytes ES120055 and ES120127 were isolated from the TCM *Saxifragaceae Astilbe* and *Cirsium shansiense* samples collected from Yunnan Province, and ES130159 were isolated from TCM *Cicuta virosa,* which was collected from Sichuan Province, China. Two desert strains LS120167 and LS120194 were collected from Taklimakan Desert. All the marine actinomycetes were isolated from sediment samples collected from the South China Sea.

Fermentation. Small scaled fermentation was conducted by culturing the strains on ISP2 agar plates at 28 °C for 7 days. A 250 mL Erlenmeyer flask pre-culture of each strain, containing 40 mL of ISP2 liquid medium, was inoculated with pieces of well-grown agar cultures of the strain at 28 °C (220 rpm) for 48 h. The pre-culture was used to inoculate in a 250 mL Erlenmeyer flask (each with 5 mL of pre-culture and 40 mL ISP2 medium) and incubated at 28 °C (220 rpm) for 3 days to obtain seed cultures for fermentation. Each of

the seed cultures was aseptically transferred to 250 mL Erlenmeyer flasks containing 40 mL of different fermentation media and was harvested for 7 days incubation at 28 °C 220 rpm; 1mL of each seed culture was also stored in 50% glycerin and water at -80℃. The broth of each strain was then transferred to a 50 mL tube and incubated at 28 °C 220 rpm for another 2 hours with the addition of 5 g HP20 resin. Each broth was centrifuged to yield supernatant and a mycelial cake with resin. The mycelial cake with resin was dried in a freeze dryer and then extracted three times with 5 mL of methanol. Crude extractions were transferred into 96 deep well plates, dried under nitrogen and stored at -20 ℃.

Large scale of Y3111, MS110104, MS110167 and LS120194 were performed by cultivation of the strains on ISP2 agar plates at 28 °C for 7 days. A 250 mL Erlenmeyer flask pre-culture of strains, containing 50 mL of fermentation medium, was inoculated and incubated with pieces of well-grown agar cultures of the strain at 28 °C (220 rpm) for 48 h. The pre-culture was used to inoculate ten 250 mL Erlenmeyer flasks (each with 5 mL of pre-culture and 100 mL fermentation medium) and incubated at 28 °C (220 rpm) for 3 days to obtain seed cultures for large- scaled fermentation. Each of the seed cultures was aseptically transferred to 80 250 mL Erlenmeyer flasks containing 40 mL of M001 medium and was harvested after 7 days (14 days for LS120194) incubation at 28 °C 220 rpm. The broth was centrifuged to yield supernatant and biomass extracts.

Media. ISP-2 (1 L): 4 g yeast extract, 10 g malt extract, 4 g dextrose, 20 g agar, pH 7.2

AM2 (1 L): soluble starch 5 g, glucose 20 g, soybean meal 10 g, peptone 2 g, yeast extract 2 g, NaCl 4 g, K_2HPO_4 0.5 g, $MgSO_4 \cdot 7H_2O$ 0.5 g, $CaCO_3$ 2 g, pH 7.8

NM2 (1L): glucose 1 g, lactose 10 g, glycerol 20 mL, soy peptone 5 g, NH_4NO_3 1.5g, yeast extract 1 g, trace element 0.2 mL/L, pH 6.0

MPG (1 L): glucose 10 g, millet powder 20 g, soybean meal 20 g, Mops 20 g, pH 7.0

M001 (1L): starch 20 g, peptone 4 g, yeast extract 8 g, CaCO3 1 g; pH 7.2 was adjusted prior to sterilization.

M12 (1 L): peptone 5 g, yeast extract 1 g, sodium citrate 0.2 g, NaCl 19.45 g, $MnCl_2$ 5.9 g, $MgSO_4$ 3.24 g, $CaCl_2$ 1.8 g, KCl 0.55 g, $NaHCO_3$ 0.5 g, KBr 0.08 g, $SrCl_2$ 0.034 g, H_3BO_3 0.022 g, $Na_2SiO_3 \cdot 9H_2O$ 0.004 g, NaF 0.002 g, NH_4NO_3 0.0016 g, Na_2PO_4 0.008 g, pH 7.2

M21 (1 L): glucose 5g, lactose 40g, soybean meal 30g, peptone 5g, K_2HPO_4 0.5g,

MgSO$_4$ 0.5g, KCl 0.3g, PH 7.0

R-1 (1 L): 20.0 g of D -mannitol, 20.0 g of D -glucose, 5 g of yeast extract, 10.0 g of peptone, 0.5 g of KH$_2$PO$_4$, 0.3 g of MgSO$_4$, 1 g of corn syrup

O-1 (1 L): 3 g of sucrose, 0.3 g of NaNO$_3$, 0.1 g of K$_2$HPO$_4$, 0.05 g of KCl, and 0.001 g of FeSO$_4$, 0.4 g MgCl$_2$

T-1 (1 L): 4.0 g yeast extract, 10.0 g malt extract and 10.0 g glucose (pH 7.0).

T-2 (1 L): glucose 60 g, yeast extract 2 g, (NH$_4$)$_2$SO$_4$ 2 g, MgSO$_4$·7H$_2$O 0.1 g, K$_2$HPO$_4$ 0.5 g, NaCl 2 g, FeSO$_4$·7H$_2$O 0.05 g, ZnSO$_4$·7H$_2$O 0.05 g, MnSO$_4$·4H$_2$O 0.05 g, CaCO$_3$ 5 g, pH 7.0

T-3 (1 L): 10 g starch, 4 g yeast extract, 2 g peptone, 1 g CaCO$_3$, 40 mg Fe$_2$(SO$_4$)$_3$·4H$_2$O, 100 mg KBr

T-4 (1 L): 24 g soluble starch, 30 g meat extract, 5g tryptose, 5 g yeast extract, 1 g glucose and 2 g calcium carbonate, pH 7.4

T-5 (1 L): 15 g soluble starch, 5 g soybean meal, 15 g peptone, 15 g glycerol, 2 g CaCO$_3$, pH 7.4

A-1 (1 L): 10 g soluble starch, 10 g yeast extract, 0.5 g KH$_2$PO$_4$, 3 g corn syrup, 20 g glucose, 5g MgSO$_4$·7H$_2$O, 3 g beef extract, 2 g CaCO$_3$, pH 7.0

A-2 (1 L): 10 g starch, 10 g glucose, 10 mL glycerol, 5 g Polypepton, 2 g yeast extract, 1 g NaCl, 3.2 g CaCO$_3$, and 1.0 mL of corn steep liquor, pH 7.4

PKS-1 (1 L): soluble starch 20.0 g; glucose 10.0 g; peptone 5.0 g; yeast extract 5.0 g; NaCl 4.0 g; K$_2$HPO$_4$ 0.5 g; MgSO$_4$·7H$_2$O, 0.5 g; CaCO$_3$ 2.0 g

PKS-2 (1 L): soybean powder 20 g, mannitol 20 g

M-1 (1 L): 10 g glucose, 10 g maltose and 3 g yeast extract, 0.15 g CaCl$_2$, 0.2 g MgCl$_2$, 30 g soybean cake meal, pH 7.2

M-2 (1 L): 24 g starch, 1 g glucose, 3 g peptone, 3 g meat extract, 5 g yeast extract, 4 g CaCO$_3$, trace metals 5 mL, pH 7.0

M-3 (1L): potatoes (sliced washed unpeeled) 200g; dextrose 20g

M-4 (1 L): 20g starch, 8.7g glycerol, 10g soybean meal, 0.01 g ZnSO$_4$·7H$_2$O, 0.005g CuSO$_4$·5H$_2$O, 0.01g FeSO$_4$·7H$_2$O

M-5 (1 L): 5 g Soybean flour, 15 g soluble starch, 15 g glycerin, 15 g peptone, 2 g CaCO$_3$, pH 7.4

E-1 (1 L): 10 g glucose, 40 g dextrin, 25 g Bactosoytone, 1 g yeast extract, 3 g $CaCO_3$, pH 7.0

D-1 (1 L): glucose 6 g; yeast extract 4 g; pH 7.0

D-2 (1 L): 2 g KH_2PO_4, 1.5 g NH_4Cl, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.5 g NaCl, 10 g glycerol, 0.4 g myoinositol, 5 g monosodium L -glutamate onohydrate, 0.084 g NaF, 0.025 g $FeSO_4 \cdot 7H_2O$, 0.01 g $ZnSO_4 \cdot 7H_2O$, 0.01 g $CoCl_2 \cdot 6H_2O$, 0.25 g $CaCO_3$, 0.001 g p-aminobenzoic acid, pH 7.0

D-3 (1 L): oatmeal 20.0 g; trace salts solution 1.0 ml; pH 7.3

B-1 (1 L): 10 g glucose, 40 g soluble amylum, 5 g yeast extract, 25 g soybean powder, 5 g peptone, 2 g $CaCO_3$, 8 g $MgSO_4 \cdot 7H_2O$, 6 g $FeSO_4 \cdot 7H_2O$, 2 g $ZnSO_4 \cdot 7H_2O$, 2 g $MnSO_4 \cdot H_2O$, 0.5 g $CoCl_2 \cdot 6H_2O$, 2 g $Na_2MoO_4 \cdot 2H_2O$, pH 7.0

Trace Salts Solution (100 mL): $FeSO_4 \cdot 7H_2O$ 0.1g; $MnCl_2 \cdot 4H_2O$ 0.1 g; $ZnSO_4 \cdot 7H_2O$ 0.1g

Anti-BCG assay. The BCG used was a *M. bovis* BCG 1173P2 strain transformed with green fluorescent protein (GFP) constitutive expression plasmid pUV3583c with direct readout of fluorescence as a measure of bacterial growth. BCG was grown at 37 °C to mid log phase in Middle brook 7H9 broth (Becton Dickinson) supplemented with 10 % OADC enrichment (Becton Dickinson) 0.05 % tween-80 and 0.2 % glycerol, which then adjusted to OD600=0.025 with culture medium as bacterial suspension. Aliquots (80 μL) of the bacterial suspension were added to each well of the 96-well microplates (clear flat-bottom), followed by adding compounds (2 μL in DMSO), which were serially twofold diluted. Isoniazid served as positive control and DMSO as negative control. The plate was incubated at 37 °C for 3 days, and GFP fluorescence was measured with Multi-label Plate Reader using the bottom read mode, with excitation at 485 nm and emission at 535 nm. MIC is defined as the minimum concentration of drug that inhibits more than 90 % of bacterial growth reflected by fluorescence value.

Antimicrobial assay. Antimicrobial assays were performed according to the Antimicrobial Susceptibility Testing Standards outlined by the Clinical and Laboratory Standards Institute (CLSI) (NCCLS 1999) using the bacteria *Staphylococcus aureus* (ATCC 6538), *Bacillus subtilis* (ATCC 6633), methicillin-resistant *S. aureus* (MRSA) and *Pseudomonas aeruginosa* (PAO1). For each organism, a loopful of glycerol stock was streaked on an LB agar plate and incubated overnight at 37 °C. A single bacterial colony

was picked and suspended in Mueller-Hinton Broth to approximately 1×10^4 CFU/mL. A twofold serial dilution of each compound to be tested (4,000 to 31.25 μg/mL in DMSO) was prepared, and an aliquot of each dilution (2 μL) was added to a 96-well flat-bottom microtiter plate (Greiner). Vancomycin and ciprofloxacin were used as positive controls and DMSO as the negative control. An aliquot (78 μL) of bacterial suspension was then added to each well (to give final compound concentrations of 100 to 0.78 μg/mL in 2.5 % DMSO), and the plate was incubated at 37 °C aerobically for 16 h. Finally, the optical density of each well at 600 nm was measured with an EnVision 2103 Multi-label Plate Reader (Perkin-Elmer Life Sciences). MIC values were defined as the minimum concentration of compound that inhibited visible bacterial growth. All the experiments were performed in triplicate.

Time-kill studies. Bacillus Calmette-Guérin Pasteur 1173P2 was cultured for 6 days at 37 °C in 7H9 broth. Cells were diluted in medium to an initial OD_{600} of 0.025 as bacterial suspension. Either compound heraclemycin C, adjusted to final concentrations of 0.5, 2, and 8 times of the MIC, or heraclemycin D, adjusted to final concentrations of 0.5, 2, and 4 times the MIC, or heraclemycins A and B, adjusted to final concentrations of 100, 25, and 6.25 μg/mL, or DMSO was then added. Aliquots (80 μL) of the cultures were removed at 0, 12, 24, 36, 48, 60, 72, 84, 96, and 108 h of incubation, and GFP fluorescence was measured with Multi-label Plate Reader using the bottom read mode, with excitation at 485 nm and emission at 535 nm. Rates of killing were determined by measuring the reduction in GFP fluorescence of viable bacteria ($\log_{10}$ RFU/ mL) at each time point with fixed concentrations of the compound. Experiments were performed in triplicate.

References

1.   Organization WH. Global Tuberculosis Report 2016.1-201 (2016)
2.   Bermudez LE, Goodman J. *Mycobacterium tuberculosis* invades and replicates within type II alveolar cells. *Infect Immun.* 64(4):1400-1406 (1996)
3.   Gupta A, Kaul A, Tsolaki AG, Kishore U, Bhakta S. *Mycobacterium tuberculosis*: Immune evasion, latency and reactivation. *Immunobiology.* 217(3):363-374 (2012)
4.   Shen Y, et al. Clinical latency and reactivation of AIDS-related mycobacterial infections. *J Virol.* 78(24):14023-14032 (2004)
5.   Vujkovic-Cvijin I, et al. Dysbiosis of the Gut Microbiota Is Associated with HIV Disease Progression and Tryptophan Catabolism. *Sci Transl Med.* 5(193) (2013)
6.   Pawlowski A, Jansson M, Skold M, Rottenberg ME, Kallenius G. Tuberculosis and HIV Co-Infection. *Plos Pathog.* 8(2) (2012)
7.   Bezuidenhout J, Roberts T, Muller L, van Helden P, Walzl G. Pleural Tuberculosis in Patients with Early HIV Infection Is Associated with Increased TNF-Alpha Expression and Necrosis in Granulomas. *Plos One.* 4(1) (2009)
8.   Lawn SD, et al. Screening for HIV-Associated Tuberculosis and Rifampicin Resistance before Antiretroviral Therapy Using the Xpert MTB/RIF Assay: A Prospective Study. *Plos Med.* 8(7) (2011)
9.   Verver S, et al. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Resp Crit Care.* 171(12):1430-1435 (2005)
10.   Kirimuhuzya C. Multi-Drug/Extensively Drug Resistant Tuberculosis (Mdr/Xdr-Tb): Renewed Global Battle Against Tuberculosis? *Understanding Tuberculosis – New Approaches to Fighting Against Drug Resistance.* 3-33 (2012)
11.   Dye C, Espinal MA, Watt CJ, Mbiaga C, Williams BG. Worldwide incidence of multidrug-resistant tuberculosis. *J Infect Dis.* 185(8):1197-1202 (2002)
12.   Bottger EC, Springer B. Tuberculosis: drug resistance, fitness, and strategies for global control. *Eur J Pediatr.* 167(2):141-148 (2008)
13.   Falzon D, et al. WHO guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update. *Eur Respir J.* 38(3):516-528 (2011)
14.   Ferdinand S, et al. Molecular characterization and drug resistance patterns of strains of *Mycobacterium tuberculosis* isolated from patients in an AIDS counseling center in Port-au-Prince, Haiti: a 1-year study. *J Clin Microbiol.* 41(2):694-702 (2003)
15.   Gonzalez-Martin J, et al. Consensus Document on the Diagnosis, Treatment and Prevention of Tuberculosis. *Arch Bronconeumol.* 46(5):255-274 (2010)
16.   Claude K. Multi-Drug/Extensively Drug Resistant Tuberculosis (Mdr/Xdr-Tb): Renewed Global Battle Against Tuberculosis? *Understanding Tuberculosis: New Approaches to Fighting Against Drug Resistance.* 3-32 (2012)
17.   Velayati AA, et al. Emergence of New Forms of Totally Drug-Resistant Tuberculosis Bacilli Super Extensively Drug-Resistant Tuberculosis or Totally Drug-Resistant Strains in Iran. *Chest.* 136(2):420-425 (2009)
18.   Rowland K. Totally drug-resistant TB emerges in India. *Nature News & Comment.* 1-4 (2012)
19.   Schraufnagel DE. Tuberculosis treatment for the beginning of the next century. *Int J Tuberc Lung D.* 3(8):651-662 (1999)
20.   P. Yadav PD, V. Kanase, S. Mishra. Overview of new anti-TB drugs. *International Journal of Pharma Sciences and Research.* 3(8):2472-2481 (2012)
21.   Grosset J. Present and new drug regimens in chemotherapy and chemoprophylaxis of tuberculosis. *Bulletin of the International Union against Tuberculosis and Lung Disease.* 65(2-3):86-91 (1990)
22.   Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. *Am J Resp Crit Care.* 167(11):1472-1477 (2003)
23.   Lawn SD, Bekker LG, Middelkoop K, Myer L, Wood R. Impact of HIV infection on the epidemiology of tuberculosis in a peri-urban community in South Africa: the need for age-

specific interventions. *Clin Infect Dis.* 42(7):1040-1047 (2006)

24. Prasad R, Verma SK, Sahai S, Kumar S, Jain A. Efficacy and safety of kanamycin, ethionamide, PAS and cycloserine in multidrug-resistant pulmonary tuberculosis patients. *The Indian journal of chest diseases & allied sciences.* 48(3):183-186 (2006)

25. Gandhi NR, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *Lancet.* 375(9728):1830-1843 (2010)

26. Mitscher LA. Bacterial topoisomerase inhibitors: quinolone and pyridone antibacterial agents. *Chemical reviews.* 105(2):559-592 (2005)

27. Rodriguez JC, Ruiz M, Climent A, Royo G. In vitro activity of four fluoroquinolones against *Mycobacterium tuberculosis.* *International journal of antimicrobial agents.* 17(3):229-231 (2001)

28. Nuermberger EL, et al. Moxifloxacin-containing regimens of reduced duration produce a stable cure in murine tuberculosis. *Am J Respir Crit Care Med.* 170(10):1131-1134 (2004)

29. Dorman SE, et al. Substitution of moxifloxacin for isoniazid during intensive phase treatment of pulmonary tuberculosis. *Am J Respir Crit Care Med.* 180(3):273-280 (2009)

30. Rivers EC, Mancera RL. New anti-tuberculosis drugs in clinical trials with novel mechanisms of action. *Drug Discov Today.* 13(23-24):1090-1098 (2008)

31. Mukherjee T, Boshoff H. Nitroimidazoles for the treatment of TB: past, present and future. *Future medicinal chemistry.* 3(11):1427-1454 (2011)

32. Singh R, et al. PA-824 Kills Nonreplicating *Mycobacterium tuberculosis* by Intracellular NO Release. *Science.* 322(5906):1392-1395 (2008)

33. Stover CK, et al. A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature.* 405(6789):962-966 (2000)

34. Matsumoto M, et al. OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice. *Plos Med.* 3(11):e466 (2006)

35. Saliu OY, Crismale C, Schwander SK, Wallis RS. Bactericidal activity of OPC-67683 against drug-tolerant *Mycobacterium tuberculosis.* *The Journal of antimicrobial chemotherapy.* 60(5):994-998 (2007)

36. Ibrahim M, et al. Synergistic activity of R207910 combined with pyrazinamide against murine tuberculosis. *Antimicrobial agents and chemotherapy.* 51(3):1011-1015 (2007)

37. Rustomjee R, et al. Early bactericidal activity and pharmacokinetics of the diarylquinoline TMC207 in treatment of pulmonary tuberculosis. *Antimicrobial agents and chemotherapy.* 52(8):2831-2835 (2008)

38. Diacon AH, et al. The diarylquinoline TMC207 for multidrug-resistant tuberculosis. *The New England journal of medicine.* 360(23):2397-2405 (2009)

39. Shinabarger D. Mechanism of action of the oxazolidinone antibacterial agents. *Expert opinion on investigational drugs.* 8(8):1195-1202 (1999)

40. Dietze R, et al. Early and extended early bactericidal activity of linezolid in pulmonary tuberculosis. *Am J Respir Crit Care Med.* 178(11):1180-1185 (2008)

41. Ntziora F, Falagas ME. Linezolid for the treatment of patients with mycobacterial infections: a systematic review. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease.* 11(6):606-611 (2007)

42. Cynamon MH, Klemens SP, Sharpe CA, Chase S. Activities of several novel oxazolidinones against *Mycobacterium tuberculosis* in a murine model. *Antimicrobial agents and chemotherapy.* 43(5):1189-1191 (1999)

43. Williams KN, et al. Promising Antituberculosis Activity of the Oxazolidinone PNU-100480 Relative to That of Linezolid in a Murine Model. *Antimicrobial agents and chemotherapy.* 53(4):1314-1319 (2009)

44. Williams KN, et al. Addition of PNU-100480 to First-Line Drugs Shortens the time Needed to Cure Murine Tuberculosis. *Am J Resp Crit Care.* 180(4):371-376 (2009)

45. Lee RE, Protopopova M, Crooks E, Slayden RA, Terrot M, Barry CE. Combinatorial lead optimization of [1,2]-diamines based on ethambutol as potential antituberculosis preclinical candidates. *J Comb Chem.* 5(2):172-187 (2003)

46. Sacksteder KA, Protopopova M, Barry CE, Andries K, Nacy CA. Discovery and development of

SQ109: a new antitubercular drug with a novel mechanism of action. *Future Microbiol.* 7(7):823-837 (2012)

47. Chen P, Gearhart J, Protopopova M, Einck L, Nacy CA. Synergistic interactions of SQ109, a new ethylene diamine, with front-line antitubercular drugs in vitro. *J Antimicrob Chemoth.* 58(2):332-337 (2006)

48. Nikonenko BV, Protopopova M, Sarnala R, Einck L, Nacy CA. Drug therapy of experimental tuberculosis (TB): Improved outcome by combining SQ109, a new diamine antibiotic, with existing TB drugs. *Antimicrobial agents and chemotherapy.* 51(4):1563-1565 (2007)

49. O'Brien RJ, Nunn PP. The need for new drugs against tuberculosis - Obstacles, opportunities, and next steps. *Am J Resp Crit Care.* 163(5):1055-1058 (2001)

50. J. Swathi KN, K. M. Sowjanya A. Krishna Satya. Marine fungal metabolites as a rich source of bioactive compounds. *African Journal of Biochemistry Research.* 7(10):184-196 (2013)

51. Juan A. Rubiolo EA, Eva Cagide. Marine Compounds as a Starting Point to Drugs. *Seafood and Freshwater Toxins.*1141-1178 (2014)

52. Damare S, Singh P, Raghukumar S. Biotechnology of marine fungi. *Progress in molecular and subcellular biology.* 53:277-297 (2012)

53. Carte BK. Biomedical potential of marine natural products. *Bioscience.* 46(4):271-286 (1996)

54. Bister B, et al. Abyssomicin C - A polycyclic antibiotic from a marine *Verrucosispora* strain as an inhibitor of the p-aminobenzoic acid/tetrahydrofolate biosynthesis pathway. *Angew Chem Int Edit.* 43(19):2574-2576 (2004)

55. Keller S, Nicholson G, Drahl C, Sorensen E, Fiedler HP, Sussmuth RD. Abyssomicins G and H and atrop-abyssomicin C from the marine *Verrucosispora* strain AB-18-032. *J Antibiot.* 60(6):391-394 (2007)

56. Wang Q, et al. Abyssomicins from the South China Sea Deep-Sea Sediment *Verrucosispora* sp.: Natural Thioether Michael Addition Adducts as Antitubercular Prodrugs. *Angew Chem Int Edit.* 52(4):1231-1234 (2013)

57. Abdalla MA, Yadav PP, Dittrich B, Schuffler A, Laatsch H. ent-Homoabyssomicins A and B, Two New Spirotetronate Metabolites from *Streptomyces* sp. Ank 210 (vol 13, pg 2156, 2011). *Org Lett.* 13(19):5409-5409 (2011)

58. Pruksakorn P, et al. Trichoderins, novel aminolipopeptides from a marine sponge-derived *Trichoderma* sp., are active against dormant mycobacteria. *Bioorg Med Chem Lett.* 20(12):3658-3663 (2010)

59. Pruksakorn P, Arai M, Liu L, Moodley P, Jacobs WR, Kobayashi M. Action-Mechanism of Trichoderin A, an Anti-dormant Mycobacterial Aminolipopeptide from Marine Sponge-Derived *Trichoderma* sp. *Biol Pharm Bull.* 34(8):1287-1290 (2011)

60. Li HX, et al. Peniphenones A-D from the Mangrove Fungus *Penicillium dipodomyicola* HN4-3A as Inhibitors of *Mycobacterium tuberculosis* Phosphatase MptpB. *J Nat Prod.* 77(4):800-806 (2014)

61. Muller D, et al. Brunsvicamides A-C: Sponge-related cyanobacterial peptides with *Mycobacterium tuberculosis* protein tyrosine phosphatase inhibitory activity. *J Med Chem.* 49(16):4871-4878 (2006)

62. Yapeng Zhanga SL, Yuchun Fanga, Tianjiao Zhua, Qianqun Gu Wei-Ming Zhu. Isolation, Structure Elucidation, and Antimycobacterial Properties of Dimeric Naphtho-g-pyrones from the Marine-Derived Fungus *Aspergillus carbonarius*. *CHEMISTRY & BIODIVERSITY.* 5:93-100 (2008)

63. Sivakumar N, Selvakumar G. Marine cyanobacteria A prolific source of antimicrobial natural products. *Antimicrobials: Synthetic and Natural Compounds.*203-231 (2016)

64. Sachiko Hasegawa AM, Masafumi Shimizu, Tomio Nishimura, Hitoshi Kunoh. Endophytic Actinomycetes and Their Interactions with Host Plants. *Actinomycetologica.* 20(2):72-81 (2006)

65. Jalgaonwala Ruby E MRT. A Review: Bacterial Endophytes and their Bioprospecting. *Journal of Pharmacy Research.* 4(3):795-799 (2011)

66. Schulz B, Boyle C, Draeger S, Rommert AK, Krohn K. Endophytic fungi: a source of novel biologically active secondary metabolites. *Mycol Res.* 106:996-1004 (2002)

67. Aly AH, Debbab A, Proksch P. Fungal endophytes: unique plant inhabitants with great

promises. *Appl Microbiol Biot.* 90(6):1829-1845 (2011)

68.   Strobel G, Daisy B. Bioprospecting for microbial endophytes and their natural products. *Microbiol Mol Biol R.* 67(4):491-+ (2003)

69.   Strobel GA. Endophytes as sources of bioactive products. *Microbes Infect.* 5(6):535-544 (2003)

70.   Tompsett R, Mcdermott W, Kidd JG. Tuberculostatic Activity of Blood and Urine from Animals Given Gliotoxin. *J Immunol.* 65(1):59-63 (1950)

71.   Intaraudom C, Boonyuen N, Suvannakad R, Rachtawee P, Pittayakhajonwut P. Penicolinates A-E from endophytic *Penicillium* sp BCC16054. *Tetrahedron Lett.* 54(8):744-748 (2013)

72.   Isaka M, Berkaew P, Intereya K, Komwijit S, Sathitkunanon T. Antiplasmodial and antiviral cyclohexadepsipeptides from the endophytic fungus *Pullularia* sp BCC 8613. *Tetrahedron.* 63(29):6855-6860 (2007)

73.   Kobayashi H, Meguro S, Yoshimoto T, Namikoshi M. Absolute structure, biosynthesis, and anti-microtubule activity of phomopsidin, isolated from a marine-derived fungus *Phomopsis* sp. *Tetrahedron.* 59(4):455-459 (2003)

74.   Isaka M, Jaturapat A, Rukseree K, Danwisetkanjana K, Tanticharoen M, Thebtaranonth Y. Phomoxanthones A and B, novel xanthone dimers from the endophytic fungus *Phomopsis* species. *J Nat Prod.* 64(8):1015-1018 (2001)

75.   Rukachaisirikul V, Sommart U, Phongpaichit S, Sakayaroj J, Kirtikara K. Metabolites from the endophytic fungus *Phomopsis* sp PSU-D15. *Phytochemistry.* 69(3):783-787 (2008)

76.   Wijeratne EMK, He HP, Franzblau SG, Hoffman AM, Gunatilaka AAL. Phomapyrrolidones A-C, Antitubercular Alkaloids from the Endophytic Fungus *Phoma* sp. NRRL 46751. *J Nat Prod.* 76(10):1860-1865 (2013)

77.   Cheng MJ, et al. Secondary metabolites from the endophytic fungus *Biscogniauxia formosana* and their antimycobacterial activity. *Phytochem Lett.* 5(3):467-472 (2012)

78.   Lam KS. Discovery of novel metabolites from marine actinomycetes. *Curr Opin Microbiol.* 9(3):245-251 (2006)

79.   Zhu H, Sandiford SK, Wezel GPV. Triggers and cues that activate antibiotic production by actinomycetes. *Journal of Industrial Microbiology.* 41(2):371-386 (2014)

80.   Wen-Jing W, Dan-Yi L, Yan-Chun L, Hui-Ming H, En-Long M, Zhan-Lin L. Caryophyllene sesquiterpenes from the marine-derived fungus Ascotricha sp. ZJ-M-5 by the one strain-many compounds strategy. *J Nat Prod.* 77(6):1367-1371 (2014)

81.   Wijeratne EMK, Paranagama PA, Gunatilaka AAL. Five new isocoumarins from Sonoran desert plant-associated fungal strains *Paraphaeosphaeria quadriseptata* and *Chaetomium chiversii.* *Tetrahedron.* 62(36):8439-8446 (2006)

82.   Powers S. In vivo and in vitro production options for fungal secondary metabolites. *Molecular Pharmaceutics.* 5(2):234-242 (2008)

83.   Ashley M. Sidebottom ARJ, Jonathan A. Karty, Darci J. Trader, Erin E. Carlson. Integrated Metabolomics Approach Facilitates Discovery of an Unpredicted Natural Product Suite from *Streptomyces coelicolor* M145. *Acs Chemical Biology.* 8(9):2009-2016 (2013)

84.   Bode HB, Bethe B, Hofs R, Zeeck A. Big effects from small changes: Possible ways to explore nature's chemical diversity. *Chembiochem.* 3(7):619-627 (2002)

85.   Doull JL, Ayer SW, Singh AK, Thibault P. Production of a novel polyketide antibiotic, jadomycin B, by *Streptomyces venezuelae* following heat shock. *The Journal of antibiotics.* 46(5):869-871 (1993)

86.   Hosaka T, et al. Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat Biotechnol.* 27(5):462-464 (2009)

87.   Recio E, Colinas A, Rumbero A, Aparicio JF, Martin JF. PI factor, a novel type quorum-sensing inducer elicits pimaricin production in *Streptomyces natalensis.* *J Biol Chem.* 279(40):41586-41593 (2004)

88.   Christian OE, Compton J, Christian KR, Mooberry SL, Valeriote FA, Crews P. Using jasplakinolide to turn on pathways that enable the isolation of new chaetoglobosins from *Phomopsis asparagi.* *J Nat Prod.* 68(11):1592-1597 (2005)

89.   Cueto M, Jensen PR, Kauffman C, Fenical W, Lobkovsky E, Clardy J. Pestalone, a new antibiotic produced by a marine fungus in response to bacterial challenge. *J Nat Prod.* 64(11):1444-

1446 (2001)

90.   Wang QX, et al. 3-Anhydro-6-hydroxy-ophiobolin A, a new sesterterpene inhibiting the growth of methicillin-resistant *Staphylococcus aureus* and inducing the cell death by apoptosis on K562, from the phytopathogenic fungus *Bipolaris oryzae*. *Bioorg Med Chem Lett.* 23(12):3547-3550 (2013)

91.   Ali K, et al. NMR spectroscopy and chemometrics as a tool for anti-TNF alpha activity screening in crude extracts of grapes and other berries. *Metabolomics.* 8(6):1148-1161 (2012)

92.   Yadav NK, Shukla P, Omer A, Pareek S, Singh RK. Next Generation Sequencing: Potential and Application in Drug Discovery. *Sci World J.* (2014)

93.   Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. *Science.* 316(5830):1497-1502 (2007)

94.   Levy S, et al. The diploid genome sequence of an individual human. *Plos Biol.* 5(10):2113-2144 (2007)

95.   Wingfield J, Wilson ID. Advances in Mass Spectrometry Within Drug Discovery. *J Biomol Screen.* 21(2):109-110 (2016)

96.   Pedro L, Quinn RJ. Native Mass Spectrometry in Fragment-Based Drug Discovery. *Molecules.* 21(8) (2016)

97.   Ashcroft AE. Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. *Nat Prod Rep.* 22(4):452-464 (2005)

98.   Cammarata MB, Thyer R, Rosenberg J, Ellington A, Brodbelt JS. Structural Characterization of Dihydrofolate Reductase Complexes by Top-Down Ultraviolet Photodissociation Mass Spectrometry. *J Am Chem Soc.* 137(28):9128-9135 (2015)

99.   Li HL, Wongkongkathep P, Van Orden SL, Loo RRO, Loo JA. Revealing Ligand Binding Sites and Quantifying Subunit Variants of Noncovalent Protein Complexes in a Single Native Top-Down FTICR MS Experiment. *J Am Soc Mass Spectr.* 25(12):2060-2068 (2014)

100.   Hida M, Satoh H, Mitsui T. Comparative study of a cluster analysis and a principal-component analysis using a polarized imaging technique for discriminating adhesive cloth tapes. *Analytical sciences : the international journal of the Japan Society for Analytical Chemistry.* 18(6):717-722 (2002)

101.   Hornquist M, Hertz J, Wahde M. Effective dimensionality for principal component analysis of time series expression data. *Bio Systems.* 71(3):311-317 (2003)

102.   Hervé Abdi, Williams LJ. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics.* 2(4):433–459 (2010)

103.   Bro R, Smilde AK. Principal component analysis. *Anal Methods-Uk.* 6(9):2812-2831 (2014)

104.   Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *Journal of pharmacological and toxicological methods.* 44(1):235-249 (2000)

105.   Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug discovery today. Technologies.* 1(4):337-341 (2004)

106.   Veber DF JS, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* 45(12):2615-2623 (2002)

107.   Tetko IV. Computing chemistry on the web. *Drug Discov Today.* 10(22):1497-1500 (2005)

108.   Tetko IV, et al. Virtual computational chemistry laboratory - design and description. *J Comput Aid Mol Des.* 19(6):453-463 (2005)

109.   Larsson J, Gottfries J, Bohlin L, Backlund A. Expanding the ChemGPS chemical space with natural products. *J Nat Prod.* 68(7):985-991 (2005)

110.   Rosen J, Gottfries J, Muresan S, Backlund A, Oprea TI. Novel Chemical Space Exploration via Natural Products. *J Med Chem.* 52(7):1953-1962 (2009)

111.   Larsson J, Gottfries J, Muresan S, Backlund A. ChemGPS-NP: Tuned for navigation in biologically relevant chemical space. *J Nat Prod.* 70(5):789-794 (2007)

112.   Rosen J, Lovgren A, Kogej T, Muresan S, Gottfries J, Backlund A. ChemGPS-NPWeb: chemical space navigation online. *J Comput Aid Mol Des.* 23(4):253-259 (2009)

113.   Dashti Y, Grkovic T, Quinn RJ. Predicting natural product value, an exploration of anti-TB drug space. *Nat Prod Rep.* 31(8):990-998 (2014)

114.   Lambert PA. Cellular impermeability and uptake of biocides and antibiotics in Gram-positive

bacteria and mycobacteria. *Journal of applied microbiology.* 92 Suppl:46S-54S (2002)

115. Ertl P, Rohde B, ., Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem.* 43(20):3714-3717 (2000)

116. Tronde A, Norden B, Marchner H, Wendel AK, Lennernas H, Bengtsson UH. Pulmonary absorption rate and bioavailability of drugs in vivo in rats: structure-absorption relationships and physicochemical profiling of inhaled drugs. *Journal of pharmaceutical sciences.* 92(6):1216-1233 (2003)

117. Andrews PR, Craik DJ, Martin JL. Functional group contributions to drug-receptor interactions. *J Med Chem.* 27(12):1648-1657 (1984)

118. Faulkner DJ. Marine natural products. *Nat Prod Rep.* 19(1):1-48 (2002)

119. Rinehart KL. Secondary Metabolites from Marine Organisms. *Ciba F Symp.* 171:236-254 (1992)

120. Fenical W. Natural-Products Chemistry in the Marine-Environment. *Science.* 215(4535):923-928 (1982)

121. Gray JS. Biomagnification in marine systems: the perspective of an ecologist. *Mar Pollut Bull.* 45(1-12):46-52 (2002)

122. Grabowski K SG. Properties and architecture of natural products revisited. *Curr Chem Biol.* 1:115-127 (2007)

123. Chomcheon P, et al. Cyclopentenones, scaffolds for organic syntheses produced by the endophytic fungus *Mitosporic dothideomycete* sp LRUB20. *J Nat Prod.* 69(9):1351-1353 (2006)

124. Pan JH, et al. Antimycobacterial activity of fusaric acid from a mangrove endophyte and its metal complexes. *Arch Pharm Res.* 34(7):1177-1181 (2011)

125. Wu MD, et al. Phytochemical investigation of *Annulohypoxylon ilanense*, an endophytic fungus derived from Cinnamomum species. *Chem Biodivers.* 10(3):493-505 (2013)

126. Verma VC, Lobkovsky E, Gange AC, Singh SK, Prakash S. Piperine production by endophytic fungus *Periconia* sp. isolated from *Piper longum* L. *The Journal of antibiotics.* 64(6):427-431 (2011)

127. Selvin J, et al. Antibacterial potential of antagonistic Streptomyces sp. isolated from marine sponge *Dendrilla nigra*. *FEMS microbiology ecology.* 50(2):117-122 (2004)

128. Ashforth EJ, et al. Bioprospecting for antituberculosis leads from microbial metabolites. *Nat Prod Rep.* 27(11):1709-1719 (2010)

129. Berdy J. Bioactive microbial metabolites - A personal view. *J Antibiot.* 58(1):1-26 (2005)

130. Bull AT, Goodfellow M, Slater JH. Biodiversity as a Source of Innovation in Biotechnology. *Annu Rev Microbiol.* 46:219-252 (1992)

131. Liu XT, et al. Systematics-guided bioprospecting for bioactive microbial natural products. *Anton Leeuw Int J G.* 101(1):55-66 (2012)

132. Dhiman RK, Schaeffer ML, Bailey AM, Testa CA, Scherman H, Crick DC. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase (IspC) from *Mycobacterium tuberculosis*: towards understanding mycobacterial resistance to fosmidomycin. *J Bacteriol.* 187(24):8395-8402 (2005)

133. Wang JF, et al. Antituberculosis Agents and an Inhibitor of the para-Aminobenzoic Acid Biosynthetic Pathway from Hydnocarpus anthelminthica Seeds. *Chemistry & Biodiversity.* 7(8):2046-2053 (2010)

134. Christensen MJ, Bennett RJ, Schmid J. Growth of Epichloe/Neotyphodium and p-endophytes in leaves of Lolium and Festuca grasses. *Mycol Res.* 106:93-106 (2002)

135. Owen NL, Hundley N. Endophytes--the chemical synthesizers inside plants. *Sci Prog.* 87(Pt 2):79-99 (2004)

136. Owen NL HN. Endophytes-the chemical synthesizers inside plants. *Science progress.* 87(2):79-99 (2004)

137. Brockmann H. Indomycin und indomycinone. *Angew Chem Int Ed Engl.* 7:493 (1968)

138. Bogdanovich T, et al. Antistaphylococcal activity of DX-619, a new des-F(6)-quinolone, compared to those of other agents. *Antimicrobial agents and chemotherapy.* 49(8):3325-3333 (2005)

139. Schumacher RW, Davidson BS, Montenegro DA, Bernan VS. Gamma-Indomycinone, a New Pluramycin Metabolite from a Deep-Sea Derived Actinomycete. *J Nat Products.* 58(4):613-617 (1995)

140. Strobel G, Daisy B, Castillo U, Harper J. Natural products from endophytic microorganisms. *J Nat Prod.* 67(2):257-268 (2004)

141. Fiedler HP, et al. Marine actinomycetes as a source of novel secondary metabolites. *Anton Leeuw Int J G.* 87(1):37-42 (2005)

142. Jensen PR, Mincer TJ, Williams PG, Fenical W. Marine actinomycete diversity and natural product discovery. *Anton Leeuw Int J G.* 87(1):43-48 (2005)

143. Subramani R, Aalbersberg W. Culturable rare Actinomycetes: diversity, isolation and marine natural product discovery. *Appl Microbiol Biot.* 97(21):9291-9321 (2013)

144. Camp D, Davis RA, Campitelli M, Ebdon J, Quinn RJ. Drug-like Properties: Guiding Principles for the Design of Natural Product Libraries. *J Nat Prod.* 75(1):72-81 (2012)

145. Kuo MS, Yurek DA, Kloosterman DA. Assignment of H-1 and C-13 Nmr Signals and the Alkene Geometry at C-7 in Borrelidin. *J Antibiot.* 42(6):1006-1007 (1989)

146. Sugawara A, et al. Borrelidin analogues with antimalarial activity: Design, synthesis and biological evaluation against Plasmodium falciparum parasites. *Bioorg Med Chem Lett.* 23(8):2302-2305 (2013)

147. Arao T, Kinjo J, Nohara T, Isobe R. Oleanene-Type Triterpene Glycosides from Puerariae Radix .2. Isolation of Saponins and the Application of Tandem Mass-Spectrometry to Their Structure Determination. *Chem Pharm Bull.* 43(7):1176-1179 (1995)

148. Beigelman L, Haeberli P, Sweedler D, Karpeisky A. Improved synthetic approaches toward 2 '-O-methyl-adenosine and guanosine and their N-acyl derivatives. *Tetrahedron.* 56(8):1047-1056 (2000)

149. Grundmann A, Li SM. Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from *Aspergillus fumigatus*. *Microbiol-Sgm.* 151:2199-2207 (2005)

150. Lo WL, Huang JC, Huang LY, Chen CY. Isolation of new aristolactam and dioxoaporphine from the leaves of *Michelia compressa* var. lanyuensis (Magnoliaceae). *Nat Prod Res.* 24(4):326-330 (2010)

151. Li PL, Jia ZJ. A new triterpene and new sesquiterpenes from the roots of *Ligularia sagitta*. *Helv Chim Acta.* 91(9):1717-1727 (2008)

152. Goto H, Terao Y, Akai S. Synthesis of Various Kinds of Isoflavones, Isoflavanes, and Biphenyl-Ketones and Their 1,1-Diphenyl-2-picrylhydrazyl Radical-Scavenging Activities. *Chem Pharm Bull.* 57(4):346-360 (2009)

153. Kim WG, Yoon TM, Kwon HJ, Suh JW. Talosins A and B: New isoflavonol glycosides with potent antifungal activity from *Kitasatospora kifunensis* MJM341 II. Physicochemical properties and structure determination. *J Antibiot.* 59(10):640-645 (2006)

154. Hu YC, MacMillan JB. A New Peptide Isolated from a Marine Derived *Streptomyces bacillaris*. *Nat Prod Commun.* 7(2):211-214 (2012)

155. Gross H. Strategies to unravel the function of orphan biosynthesis pathways: recent examples and future prospects. *Appl Microbiol Biot.* 75(2):267-277 (2007)

156. Omura S, et al. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *P Natl Acad Sci USA.* 98(21):12215-12220 (2001)

157. Wilkinson B, Micklefield J. Mining and engineering natural-product biosynthetic pathways. *Nat Chem Biol.* 3(7):379-386 (2007)

158. Galagan JE, et al. Sequencing of *Aspergillus nidulans* and comparative analysis with A-fumigatus and A-oryzae. *Nature.* 438(7071):1105-1115 (2005)

159. Scherlach K, Hertweck C. Triggering cryptic natural product biosynthesis in microorganisms. *Org Biomol Chem.* 7(9):1753-1760 (2009)

160. Lew W, Pai M, Oxlade O, Martin D, Menzies D. Initial drug resistance and tuberculosis treatment outcomes: Systematic review and meta-analysis. *Ann Intern Med.* 149(2):123-134 (2008)

161. Pettit RK. Culturability and Secondary Metabolite Diversity of Extreme Microbes: Expanding

Contribution of Deep Sea and Deep-Sea Vent Microbes to Natural Product Discovery. *Mar Biotechnol.* 13(1):1-11 (2010)

162. Wilson ZE, Brimble MA. Molecules derived from the extremes of life. *Nat Prod Rep.* 26(1):44-71 (2009)

163. David C, Davis RA, Marc C, James E, Quinn RJ. Drug-like properties: guiding principles for the design of natural product libraries. *J Nat Prod.* 75(1):72-81 (2012)

164. Paranagama PA, Wijeratne EMK, Gunatilaka AAL. Uncovering biosynthetic potential of plant-associated fungi: Effect of culture conditions on metabolite production by *Paraphaeosphaeria quadriseptata* and *Chaetomium chiversii. J Nat Prod.* 70(12):1939-1945 (2007)

165. Yoon V, Nodwell JR. Activating secondary metabolism with stress and chemicals. *J Ind Microbiol Biot.* 41(2):415-424 (2014)

166. Puglia AM, Vohradsky J, Thompson CJ. Developmental Control of the Heat-Shock Stress Regulon in *Streptomyces Coelicolor. Mol Microbiol.* 17(4):737-746 (1995)

167. Servant P, Mazodier P. Negative regulation of the heat shock response in *Streptomyces. Arch Microbiol.* 176(4):237-242 (2001)

168. Hofs R, Walker M, Zeeck A. Hexacyclinic acid, a polyketide from *Streptomyces* with a novel carbon skeleton. *Angew Chem Int Edit.* 39(18):3258-+ (2000)

169. Wang FZ, Wei HJ, Zhu TJ, Li DH, Lin ZJ, Gu QQ. Three New Cytochalasins from the Marine-Derived Fungus *Spicaria elegans* KLA03 by Supplementing the Cultures with L- and D-Tryptophan. *Chemistry & Biodiversity.* 8(5):887-894 (2011)

170. Doull JL, Singh AK, Hoare M, Ayer SW. Conditions for the Production of Jadomycin-B by *Streptomyces-Venezuelae* Isp5230 - Effects of Heat-Shock, Ethanol Treatment and Phage Infection. *J Ind Microbiol.* 13(2):120-125 (1994)

171. Ayer SW, et al. Jadomycin, a Novel 8h-Benz[B]Oxazolo[3,2-F]Phenanthridine Antibiotic from *Streptomyces-Venezuelae* Isp5230. *Tetrahedron Lett.* 32(44):6301-6304 (1991)

172. Jakeman DL, Graham CL, Young W, Vining LC. Culture conditions improving the production of jadomycin B. *J Ind Microbiol Biot.* 33(9):767-772 (2006)

173. Hayes A, Hobbs G, Smith CP, Oliver SG, Butler PR. Environmental signals triggering methylenomycin production by *Streptomyces coelicolor* A3(2). *J Bacteriol.* 179(17):5511-5515 (1997)

174. Wang R, et al. Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in *Streptomyces coelicolor. Mol Microbiol.* 87(1):30-48 (2013)

175. Sébastien Rigali FT, Sharief Barends, Suzanne Mulder, Andreas W Thomae, David A Hopwood, Gilles P van Wezel. Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces. Scientific reports.* 9(7):670-675 (2008)

176. Makitrynskyy R, et al. Pleiotropic regulatory genes bldA, adpA and absB are implicated in production of phosphoglycolipid antibiotic moenomycin. *Open Biol.* 3(10) (2013)

177. Corre C, Song LJ, O'Rourke S, Chater KF, Challis GL. 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *P Natl Acad Sci USA.* 105(45):17510-17515 (2008)

178. Boulahya KA, et al. OdhI dephosphorylation kinetics during different glutamate production processes involving Corynebacterium glutamicum. *Appl Microbiol Biot.* 87(5):1867-1874 (2010)

179. Zhang CH, Fevereiro PS. The effect of heat shock on paclitaxel production in *Taxus yunnanensis* cell suspension cultures: Role of abscisic acid pretreatment. *Biotechnol Bioeng.* 96(3):506-514 (2007)

180. Rao YM, Sureshkumar GK. Improvement in bioreactor productivities using free radicals: HOCl-induced overproduction of xanthan gum from *Xanthomonas campestris* and its mechanism. *Biotechnol Bioeng.* 72(1):62-68 (2001)

181. Dash S, Mohanty N. Response of seedlings to heat-stress in cultivars of wheat: Growth temperature-dependent differential modulation of photosystem 1 and 2 activity, and foliar antioxidant defense capacity. *J Plant Physiol.* 159(1):49-59 (2002)

182. Chatterjee I SG, Heilmann C, Sahl HG, Maurer HH, Herrmann M. Very low ethanol concentrations affect the viability and growth recovery in post-stationary-phase *Staphylococcus aureus* populatons. *Appl. Environ. Microbiol. .* 72:2627-2636 (2006)

183. Zhou WW, Ma B, Tang YJ, Zhong JJ, Zheng XD. Enhancement of validamycin A production by addition of ethanol in fermentation of *Streptomyces hygroscopicus* 5008. *Bioresource Technol.* 114:616-621 (2012)

184. Scheidt KA, Tasaka A, Bannister TD, Wendt MD, Roush WR. Total synthesis of bafilomycin A1. *Abstr Pap Am Chem S.* 216:U502-U502 (1998)

185. Carr G, et al. Bafilomycins Produced in Culture by *Streptomyces* spp. Isolated from Marine Habitats Are Potent Inhibitors of Autophagy. *J Nat Prod.* 73(3):422-427 (2010)

186. Axel Kretschmer MD, Martin Deeg & Hanspaul Hagenmaier. The Structures of Novel Insecticidal Macrolides: Bafilomycins D and E, and Oxohygrolidin. *Agricultural and Biological Chemistry.* 49(8):2509-2511 (1985)

187. Yu ZG, et al. Bafilomycins produced by an endophytic actinomycete *Streptomyces* sp YIM56209. *J Antibiot.* 64(1):159-162 (2011)

188. Wilton JH, Hokanson GC, French JC. Pd-118,576 - a New Antitumor Macrolide Antibiotic. *J Antibiot.* 38(11):1449-1452 (1985)

189. Lu CH, Shen YM. A new macrolide antibiotic with antitumor activity produced by *Streptomyces* sp CS, a commensal microbe of Maytenus hookeri. *J Antibiot.* 56(4):415-418 (2003)

190. Ding N, et al. Bafilomycins and Odoriferous Sesquiterpenoids from *Streptomyces albolongus* Isolated from *Elephas maximus* Feces. *J Nat Prod.* 79(4):799-805 (2016)

191. Frandberg E, Petersson C, Lundgren LN, Schnurer J. *Streptomyces halstedii* K122 produces the antifungal compounds bafilomycin B1 and C1. *Can J Microbiol.* 46(8):753-758 (2000)

192. Goetz MA, et al. L-155,175 - a New Antiparasitic Macrolide Fermentation, Isolation and Structure. *J Antibiot.* 38(2):161-168 (1985)

193. Vanek Z, Mateju J, Curdova E. Immunomodulators Isolated from Microorganisms. *Folia Microbiol.* 36(2):99-111 (1991)

194. Bowman EJ, Siebers A, Altendorf K. Bafilomycins - a Class of Inhibitors of Membrane Atpases from Microorganisms, Animal-Cells, and Plant-Cells. *P Natl Acad Sci USA.* 85(21):7972-7976 (1988)

195. Graupner PR, et al. Dihydromaltophilin; A novel fungicidal tetramic acid containing metabolite from *Streptomyces* sp. *J Antibiot.* 50(12):1014-1019 (1997)

196. Li YY, Huffman J, Li Y, Du LC, Shen YM. 3-Hydroxylation of the polycyclic tetramate macrolactam in the biosynthesis of antifungal HSAF from *Lysobacter enzymogenes* C3. *Medchemcomm.* 3(8):982-986 (2012)

197. Kazuuoshi YK, Mutsuya A, Heiichi S. A new antibiotic, Ikarugamycin. *The Journal of antibiotics.* 25(5):271-280 (1972)

198. Hashidoko Y, Nakayama T, Homma Y, Tahara S. Structure elucidation of xanthobaccin A, a new antibiotic produced from *Stenotrophomonas* sp strain SB-K88. *Tetrahedron Lett.* 40(15):2957-2960 (1999)

199. Nakayama T, Homma Y, Hashidoko Y, Mizutani J, Tahara S. Possible role of xanthobaccins produced by *Stenotrophomonas* sp strain SB-K88 in suppression of sugar beet damping-off disease. *Appl Environ Microb.* 65(10):4334-4339 (1999)

200. Miyao H, Sakai Y, Takeshita T, Kinjo J, Nohara T. Triterpene saponins from *Abrus cantoniensis* (Leguminosae) .1. Isolation and characterization of four new saponins and a new sapogenol. *Chem Pharm Bull.* 44(6):1222-1227 (1996)

201. Kubo T, et al. Study on the Constituents of *Desmodium styracifolium* .14. *Chem Pharm Bull.* 37(8):2229-2231 (1989)

202. Schobert R, Schlenk A. Tetramic and tetronic acids: An update on new derivatives and biological aspects. *Bioorgan Med Chem.* 16(8):4203-4221 (2008)

203. Jakobi M, et al. Maltophilin: A new antifungal compound produced by *Stenotrophomonas maltophilia* R3089. *J Antibiot.* 49(11):1101-1104 (1996)

204. Liu FQ, Guoliang; Shen, Yan. Antifungal and anti-oomycetes metabolite of *Lysobacter enzymogenes* OH11. *Faming Zhuanli Shenqing.* CN 104592239 A 20150506 (2015)

205. Li S, et al. An antibiotic complex from *Lysobacter enzymogenes* strain C3: Antimicrobial activity and role in plant disease control. *Phytopathology.* 98(6):695-701 (2008)

206. Li JY, Kuang Y, Shi JF, Gao YA, Zhou J, Xu B. The conjugation of nonsteroidal anti-inflammatory

drugs (NSAID) to small peptides for generating multifunctional supramolecular nanofibers/hydrogels. *Beilstein J Org Chem.* 9:908-917 (2013)

207. Wells G, et al. Design, synthesis, and biophysical and biological evaluation of a series of Pyrrolobenzodiazepine - Poly(N-methylpyrrole) conjugates. *J Med Chem.* 49(18):5442-5461 (2006)

208. Kamal A, Kumar PP, Sreekanth K, Seshadri BN, Ramulu P. Synthesis of new benzimidazole linked pyrrolo[2,1-c][1,4]benzodiazepine conjugates with efficient DNA-binding affinity and potent cytotoxicity. *Bioorg Med Chem Lett.* 18(8):2594-2598 (2008)

209. Petrusek RL, Uhlenhopp EL, Duteau N, Hurley LH. Reaction of Anthramycin with DNA - Biological Consequences of DNA Damage in Normal and Xeroderma Pigmentosum Cell-Lines. *J Biol Chem.* 257(11):6207-6216 (1982)
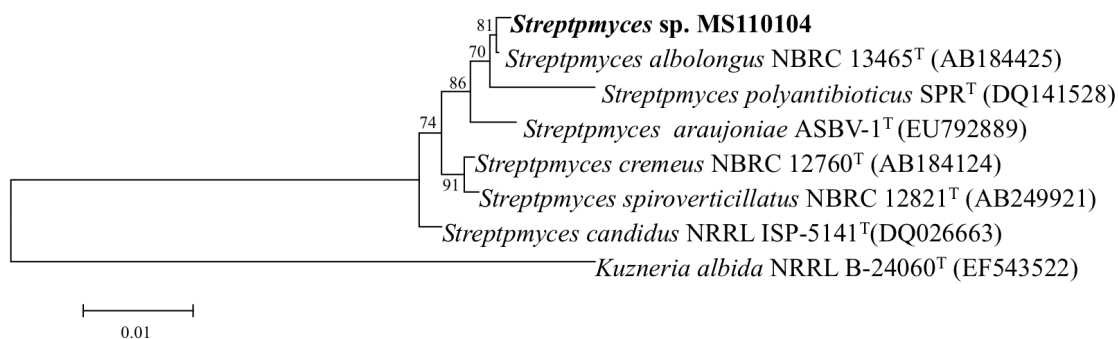
Supporting Information

Figure S1 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110104 and related species constructed by MEGA.

Figure S2 Heat map of the anti-BCG screening on 320 LLE fractions of MS110104.

Figure S3 PCA results of 320 fractions from MS110104.

Figure S4 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110167 and related species constructed by MEGA.

Figure S5 Heat map of the anti-BCG screening on 320 LLE fractions of MS110167.

Figure S6 PCA results of 320 fractions from MS110167.

Figure S7 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of LS120194 and related species constructed by MEGA.

Figure S8 Heat map of the anti-BCG screening on 320 LLE fractions of LS120194.

Figure S9 PCA results of 320 fractions from LS120194.

Figure S1 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110104 and related species constructed by MEGA.
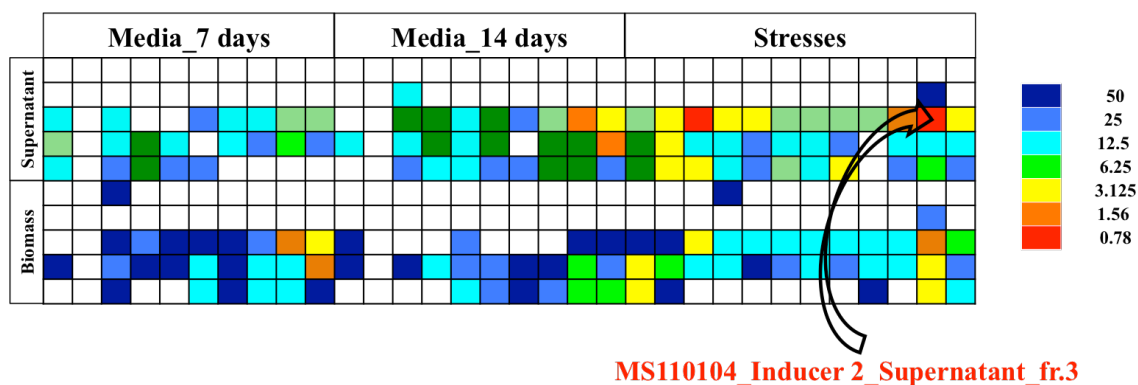

Figure S2 Heat map of the anti-BCG screening on 320 LLE fractions of MS110104. Each sample was presented in black blank.
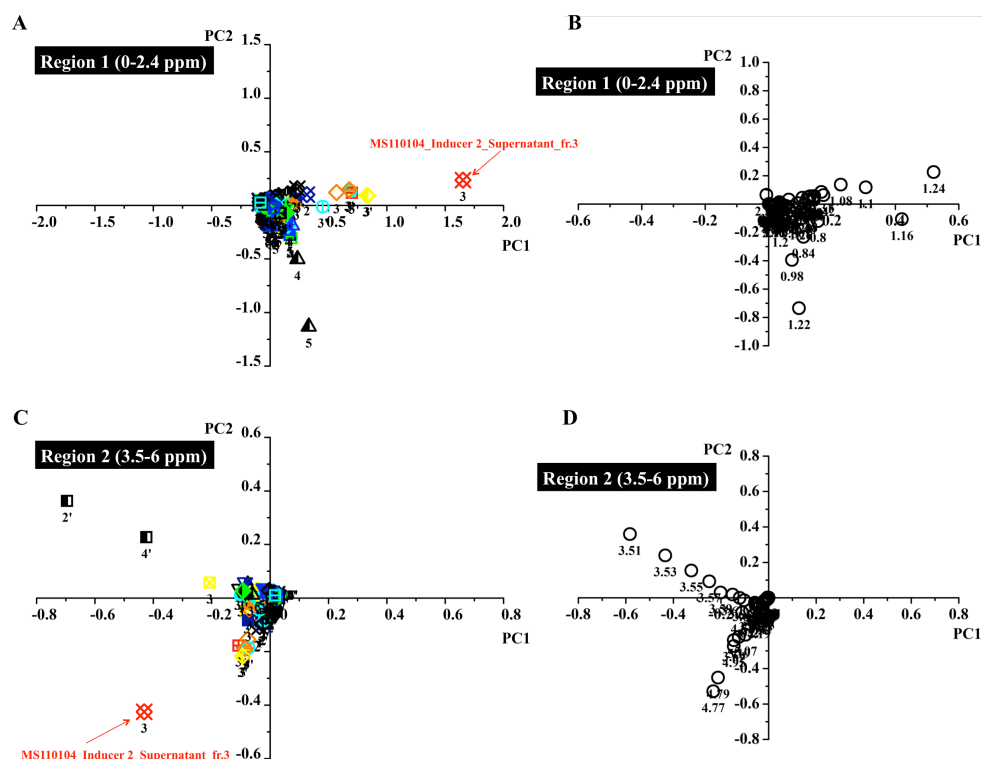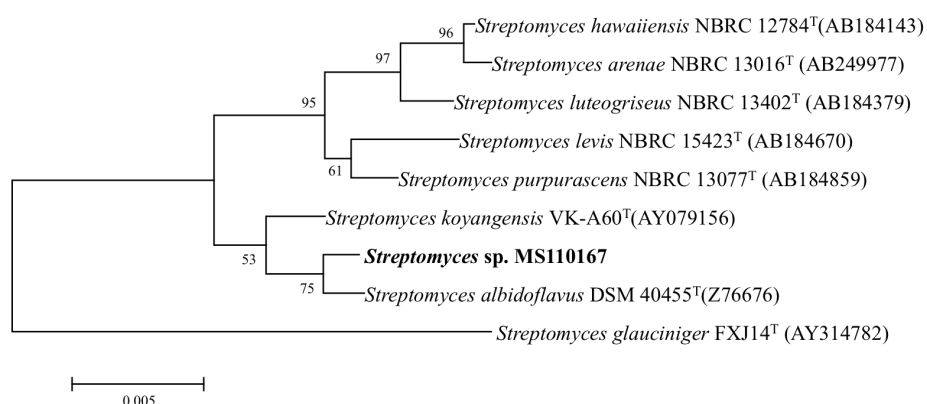
Figure S3 PCA results of 320 fractions from MS110104. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 - 2.4 ppm). Fraction numbers were shown below B. Loading plot of region 1 (0 - 2.4 ppm). C. Score plot of region 1 (3.5 - 6 ppm). Fraction numbers were shown below. D. Loading plot of region 1 (3.5 - 6 ppm). Bucket values in loading plots were shown as numbers below each circle.

Figure S4 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of MS110167 and related species constructed by MEGA.
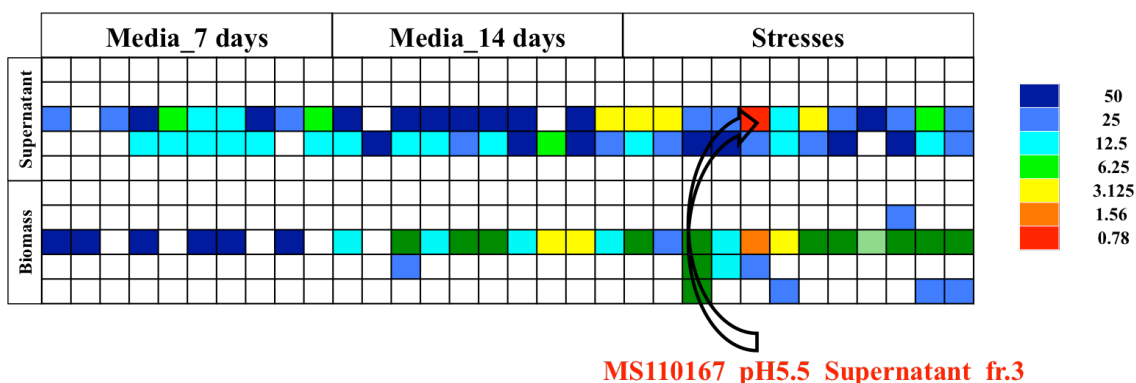
Figure S5 Heat map of the anti-BCG screening on 320 LLE fractions of MS110167. Each sample was presented in black blank.
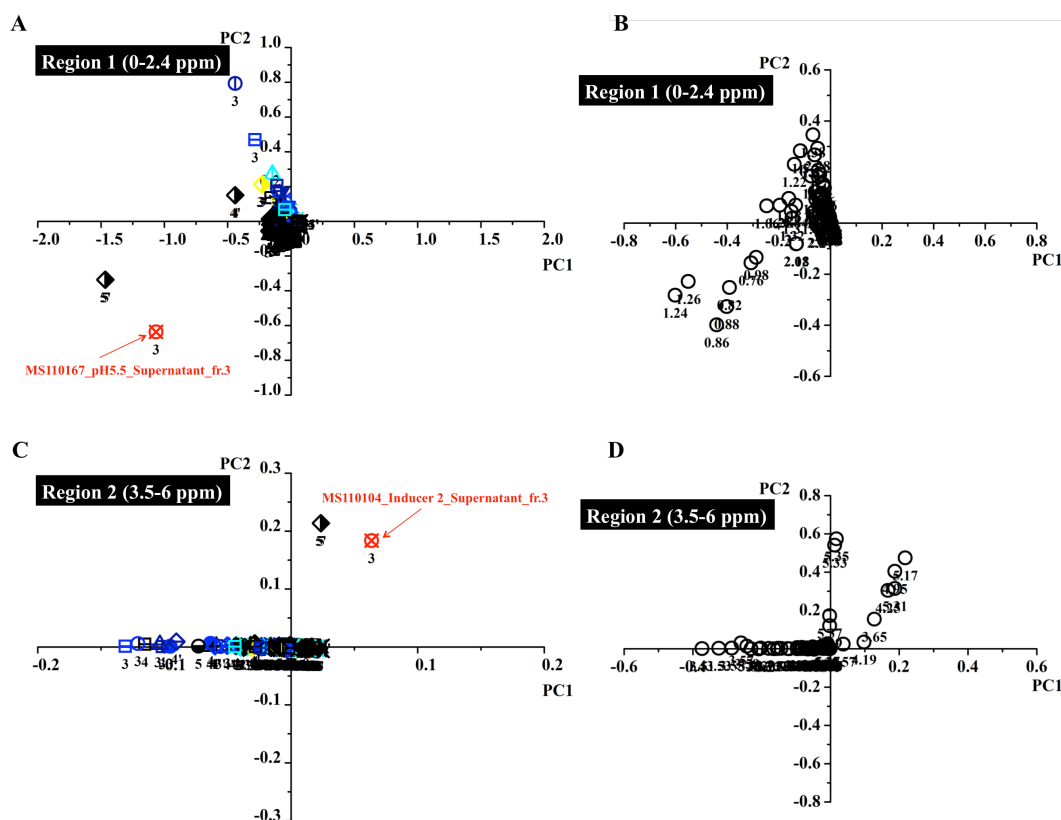

Figure S6 PCA results of 320 fractions from MS110167. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 1 (0 - 2.4 ppm). Fraction numbers were shown below B. Loading plot of region 1 (0 - 2.4 ppm). C. Score plot of region 1 (3.5 - 6 ppm). Fraction numbers were shown below. D. Loading plot of region 1 (3.5 - 6 ppm). Bucket values in loading plots were shown as numbers below each circle.
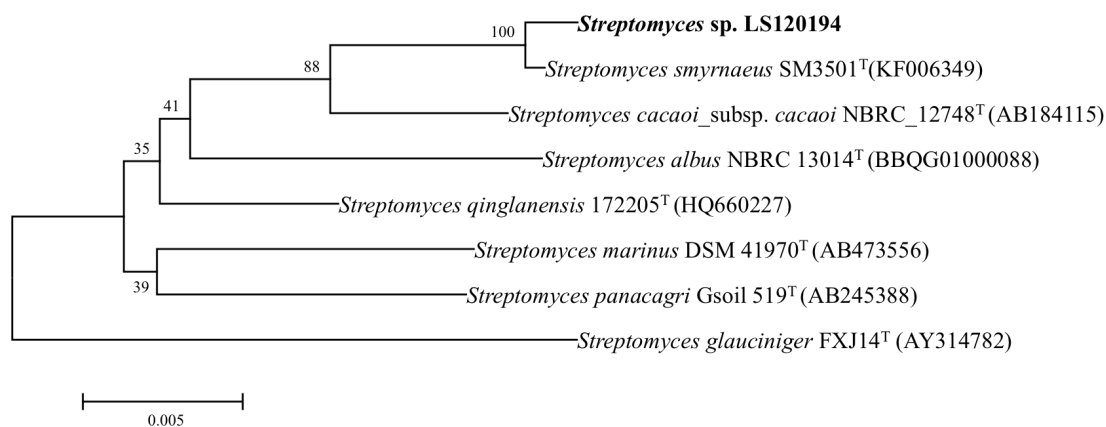
Figure S7 Neighbor-joining phylogenetic tree from the 16S rDNA sequences of LS120194 and related species constructed by MEGA.
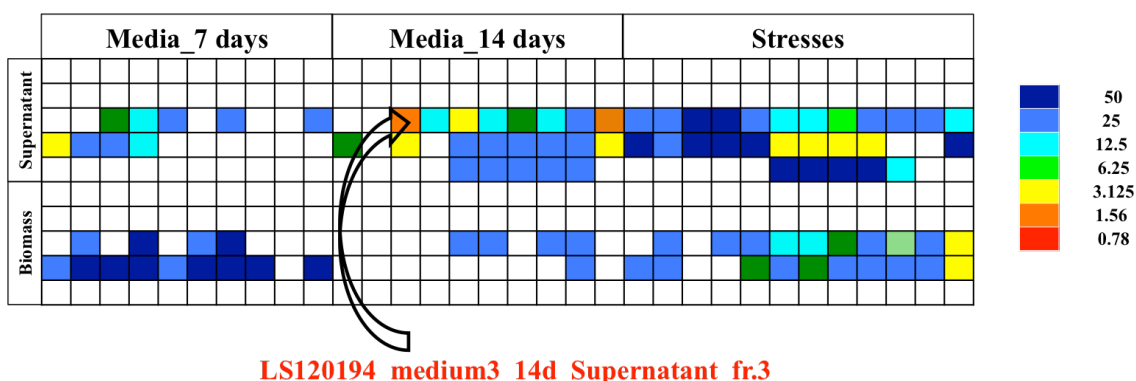

LS120194_medium3_14d_Supernatant_fr.3

Figure S8 Heat map of the anti-BCG screening on 320 LLE fractions of LS120194. Each sample was presented in a black blank.
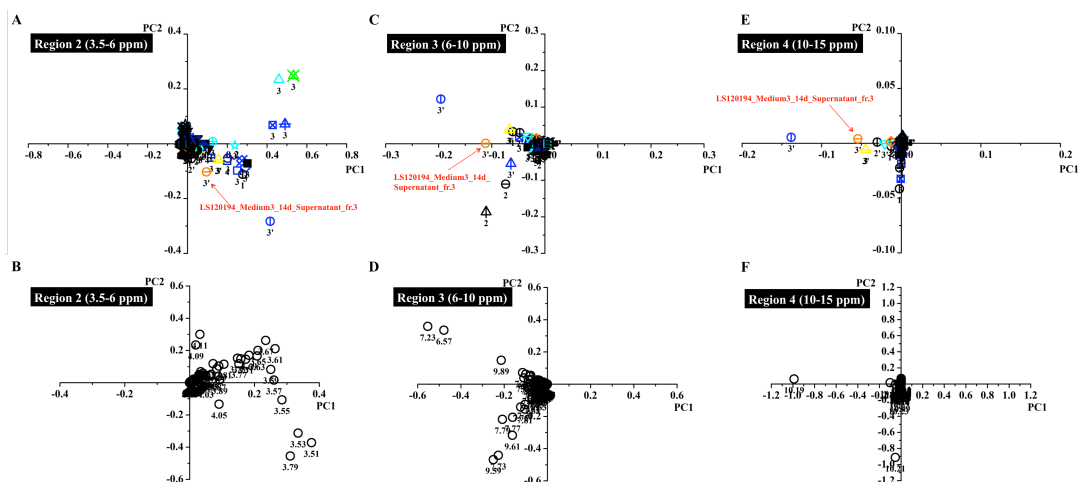
Figure S9 PCA results of 320 fractions from LS120194. Anti-BCG activity was shown in different colors coded according to the legend in Figure 6.24 (blue lowest and red highest activity). A. Score plot of region 2 (3.5 - 6 ppm). Fraction numbers were shown below B. Loading plot of region 2 (3.5 - 6 ppm). C. Score plot of region 3 (6 - 10 ppm). Fraction numbers were shown below. D. Loading plot of region 3 (6 - 10 ppm). E. Score plot of region 4 (10 - 15 ppm). Fraction numbers were shown below. D. Loading plot of region 4 (10 - 15 ppm). Bucket values in loading plots were shown as numbers below each circle.