Title

Discovery and manipulation of a mosquito-specific toxin for insecticide development

Author

Vian, Kaitlyn Nicole

Publication Date

2023

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

DISCOVERY AND MANIPULATION OF A MOSQUITO-SPECIFIC TOXIN FOR INSECTICIDE DEVELOPMENT

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

DOCTOR OF PHILOSOPHY

in

CHEMISTRY & BIOCHEMISTRY

by

Kaitlyn N. Vian

December 2023

The Dissertation of Kaitlyn N. Vian is approved:

Professor John B. MacMillan, Advisor

Professor Theodore Holman, Chair

Professor Carrie L. Partch

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents iii

List of Figures vi

List of Tables ix

Abstract x

Acknowledgments xii

Chapter 1: The need for a mosquito-specific insecticide. 1

1.1 Vector-borne diseases 2

1.2 Mosquito-borne diseases 4

1.2.1 Malaria 4

1.2.2 Dengue 7

1.3 Favorable climates for mosquitos 7

1.4 Vector control method for mosquitoes 8

1.4.1 Physical vector control 10

1.4.2 Biological vector control 10

1.4.3 Insecticide use for chemical vector control 11

1.5 Mosquito resistance of insecticides 14

1.6 Detrimental effects of insecticides 15

1.7 Summary 16

1.8 References 18

Chapter 2: Discovery of a mosquito-specific toxin NP-34 26

2.1 Introduction 27

2.1.1 Natural products in drug discovery 27

2.1.2 High-throughput screening to determine a mosquito-specific toxin 29

2.1.3 Characterization of NP-34 32

2.1.4 Known boronated natural products 33

2.1.5 NP-34 toxicity in larva and adult mosquitos 36

2.1.6 In vivo acute toxicity testing of NP-34 indicated no effect on mice 38

2.1.7 Polyketide synthases 39

2.1.8 Determining the biosynthetic gene cluster of NP-34 43

2.1.8a Initial approaches to validating the BGC of NP-34 43

2.1.8b Heterologous expression of NP-34 in *Streptomyces albus J1074* 44

2.2 Results 45

2.2.1 Confirmation of NP-34 in SNC-034 crude extracts 45

2.2.2 Sequencing and secondary metabolite analysis 48

2.2.3 Proposed biosynthetic gene cluster of NP-34 50

2.2.4 Heterologous expression of NP-34 in *Streptomyces albus J1074* 53

2.3 Discussion 56

2.4 References 57

Chapter 3: The effect of gene knockouts on the biosynthetic gene cluster of NP-34 64

3.1 Introduction 65

3.1.1 Gene Knockouts 65

3.1.2 Design of NP34 knockouts by homologous recombination 66

3.1.3 Homologous recombination with pREDCas9 69

3.2 Results 72

3.2.1 Analysis of analogs produced by *S. albus* expressing NP-34 72

3.2.2 Downstream acyltransferase controls production of NP-34 analogs 75

3.2.3 Thioesterase knockout inhibits dimerization of NP-34 77

3.2.4 LuxR is required for NP-34 production 78

3.3 Discussion 80

3.4 References 82

Chapter 4: Exploring the promiscuity of the downstream acyltransferase 86

4.1 Introduction 87

4.1.1 Initial in vitro experiments with the downstream acyltransferase 87

4.1.2 Using N-acetyl-cysteamine thioesters in place of coenzyme A 88

4.1.3 Synthesis of SNAK

4.2 Results

4.2.1 Synthesis of the acyl-CoA mimic SNAK with an alkyne handle

4.2.2 Feeding studies with SNAK

4.3 Discussion

4.4 References

Materials and Methods

List of Figures

Figure 1.1. The life cycle of an *Anopheles* mosquito

Figure 1.2. Insecticide classes and a molecular structure from each class used against mosquitos

Figure 2.1. Natural products used as insecticides

Figure 2.2. Natural product NP-34 maintains toxicity and selectivity towards mosquito cells

Figure 2.3. Safety of NP-34 against human epithelial and neuroendocrine cell lines

Figure 2.4. Boronated polyketide macrolide natural products

Figure 2.5. Acyl group in NP-34 is required for activity against mosquito cells

Figure 2.6. Toxicity and selectivity of NP-34 in adult mosquitos

Figure 2.7. NP-34 maintains toxicity in pyrethroid-resistant adult mosquitos and larvae

Figure 2.8. NP-34 shows no toxic effects in mice

Figure 2.9. Domains in each module of type I polyketide synthases with malonyl-CoA

Figure 2.10. ^{11}B-NMR spectrum of purified NP-34 and crude organic extract from SNC-034

Figure 2.11. HRMS of purified NP-34 and crude organic extracts from SNC-034 in methanol

Figure 2.12. Molecular structures and associated m/z value of NP-34 analogs observed in SNC-034 extracts by HRMS in negative mode

Figure 2.13. antiSMASH secondary metabolite BGC prediction

Figure 2.14. Comparison of antiSMASH prediction of the BGC at Region 6 of a transAT-PKS in SNC-034 and MiBIG analysis of the published tartrolon BGC

Figure 2.15. Proposed biosynthetic gene cluster of NP-34 produced by *Streptomyces malachitospinus*

Figure 2.16. Predicted AlphaFold structures of KS domains

Figure 2.17. ^{11}B-NMR of crude organic extracts from SNC-034 and *S. albus* J1074 expressing NP-34

Figure 2.18. HRMS comparison of NP-34 and analogs in SNC-034 and the NP-34 BAC expressed in *S. albus J1074*

Figure 3.1. Targeted knockouts in NP-34 gene cluster

Figure 3.2. pREDCas9 plasmid utilized during homologous recombination

Figure 3.3. Knockouts produced by homologous recombination

Figure 3.4. HRMS spectra of the proposed BGC of NP-34 BGC expressed in *S. albus*

Figure 3.5. Molecular structures of the proposed analogs produced by the BGC of NP-34

Figure 3.6. HRMS spectra of analogs produced by the proposed BGC of NP-34 produced in *S. albus*

Figure 3.7. ^{11}B-NMR and HRMS of NP-34 BGC with acyltransferase knockout heterologously expressed in *S. albus*

Figure 3.8. HRMS and ^{11}B-NMR of the thioesterase knockout in the NP-34 BAC expressed in *S. albus*

Figure 3.9. HRMS and ^{11}B-NMR of the LuxR knockout in the NP-34 BAC expressed in *S. albus*

Figure 4.1. Comparison of acyl-CoA versus acyl-SNAC

Figure 4.2. Synthesis of SNAK

Figure 4.3. Feeding study with SNAK.

Figure 4.4. ^{1}H-NMR of SNAK in $CDCl_3$

Figure 4.5. HRMS spectra of SNAK in MeOH in positive mode

Figure 4.6. HRMS spectra of organic extracts of SNAK feeding study in *S. albus* expressing NP-34 BAC in negative mode

Figure 4.7. Zoomed in HRMS spectra of organic extracts of SNAK feeding study in *S. albus* expressing NP-34 BAC in negative mode

List of Tables

Table 1. Common vector-borne diseases

Table 2. Chemical shift of ^{1}H-NMR of SNAK in $CDCl_3$

Abstract

Discovery and manipulation of a mosquito-specific toxin for insecticide development

By Kaitlyn N. Vian

Mosquito-borne diseases are the deadliest of all vector-borne diseases. Malaria alone accounted for 249 million cases and 608,000 deaths in 2022, while Dengue is estimated to put half the world's population at risk for contracting the virus. Vector control by the use of insecticides significantly reduces the risk of contracting these diseases, but harms other animals and humans in the process. Many of these insecticides are also used for other insects in agriculture. Because of this, resistance is on the rise for these non-specific insecticides. Now more than ever, a mosquito-specific toxin is needed to prevent the spread of these deadly diseases without harming other organisms. A natural product NP-34, produced by SNC-034 (*Streptomyces malachitospinus*), was found to selectively kill mosquito cells vs. *Drosophila* and *Spodoptera* cells. Further animal studies proved NP-34 was active against both adult mosquitos and larvae. Characterization of NP-34 showed a boronated polyketide macrolide from the aplasmomycin family. Sequencing and genome mining led to the discovery of the biosynthetic gene cluster (BGC) of NP-34, a trans-AT-PKS. By constructing a bacterial artificial chromosome with the NP-34 BGC, NP-34 was heterologously expressed in

Streptomyces albus J1074. Knockout of the downstream acyltransferase by homologous recombination led to the discovery that the acyltransferase controls the last step of NP-34 production. Because NP-34 has multiple analogs, this step is responsible for creating the different analogs by incorporating different acyl-groups. Knockout of the last thiosetrase (TE) led to no production of NP-34, indicating the TE could be responsible for releasing the polyketide chain off the phosphopantetheine arm and the dimerization and cyclization of the molecule. Loss of production of NP-34 after knock out of the transcriptional regulator, LuxR, could indicate LuxR is a transcriptional activator. Feeding studies with CoA mimic SNAC, indicated the possibility that new analogs of NP-34 could be produced. With this in mind, NP-34 could be further manipulated and tested to create a more selective and effective mosquitocide for the prevention of mosquito-borne diseases.

Acknowledgments

The people who have helped me the most are my family. To my entire family, thank you for your love and support! I look forward to every holiday and get together because of how amazing every one of you are. I could not have asked for a better family. To my parents Christine and Paul, you have always been supportive of everything I do and I would not have accomplished this without you. You have given me the drive to be the best person I can be. You are amazing people and I feel so fortunate to call you my parents.

To my boyfriend Jon, there are no words to describe how grateful I am for all your support. You have been there every step of the way through the good and the bad. I will never forget how you stayed up until 2:30 in the morning helping me edit my chapters and my figures. You are an amazing scientist and even better partner. I am so lucky to have you in my life.

To my fellow grad students, this has been one hell of a journey. Thank you for keeping me sane through some rough times. To my cohort, you are the best cohort this department has ever seen. I would not have stayed in this program without you all. Through strikes, covid, wildfires, floods, and power outages it's been a roller coaster of a journey.

To my friends from college, thanks for giving me the confidence to pursue grad school. To Susette, you are basically a sister to me and someone who has been there for me always. We have both changed so much the last

decade but no matter what I know you will always be there for me. I can't wait to become an Auntie!

To my PI, thank you for you support and allowing me work in your lab. This project has been unbelievably difficult, but also so rewarding. The challenges I have overcome have made me a better scientist today. To my dissertation committee Carrie and Ted, you have always been so supportive, thank you for not giving up on me! Carrie, I will never forget how you supported my dreams to go to grad school when I was an undergrad and wrote me a letter of recommendation for grad school. I still remember when you called me to tell me I was accepted into grad school. It meant so much coming from an amazing scientist like you.

To the Partch lab, you feel like a second lab family. You have always been there to help me find a reagent or talk about cloning and protein expression. The same goes for the McKinnie lab, our neighbors. It has been great to have another lab who also works with natural products. Thank you Shaun for always answering my questions and being so supportive and for your kindness and insight.

Finally to my lab, thank you for all your help. To Leah, Jocelyn, Riley, and Kevin thank you for answering all my organic chem questions and helping me with NMR, synthesis, and purification. To Rebecca, the other biochemist in our lab, thanks for letting me bounce ideas off you and helping me when I was sick or couldn't be in lab; and to Goose who always made any

day better! To Sahar, you have always been so kind, thank you. To Victor, thanks for supporting me not only in the lab but also outside of lab. You always brought so much energy into lab and it wasn't the same after you left. I can't wait to see what everyone does next, I know it will be great!

Chapter 1

The need for a mosquito-specific insecticide

1.1 Vector-borne diseases

According to the world health organization (WHO), vector-borne diseases account for more than 17 % of all infectious diseases and more than 700,000 deaths annually.[1] A vector-borne disease is caused by an infection from a parasite, bacteria, or virus spread by a vector such as fleas, ticks, flies, aquatic snails, and mosquitos. The vector, which is most likely an insect that takes blood meal, sucks the blood of an infected host such as an animal or human, and infects itself with the disease. [1,2,3] The disease is then spread to a new host when the insect takes another blood meal. Common vector-borne diseases mentioned in Table 1.1 include malaria, West Nile Virus, Zika, and Dengue spread by mosquitos, Lyme disease and encephalitis from ticks, Sleeping-sickness from Tsetse flies, Typhus from Lice, and the Plague from fleas (transmitted from rats to humans). Vector-borne diseases are mostly spread in tropical and sub-tropical climates and disproportionately affect the poorest communities.[1] Due to climate change and international travel, vector-borne diseases have emerged and reemerged in areas that were once eradicated.[3–5]

Vector	Disease caused	Type of pathogen
Aedes Mosquito	Chikungunya	Virus
	Dengue	Virus
	Lymphatic filariasis	Parasite
	Rift Valley fever	Virus
	Yellow Fever	Virus
	Zika	Virus
Anopheles Mosquito	Lymphatic filariasis	Parasite
	Malaria	Parasite
Culex Mosquito	Japanese encephalitis	Virus
	Lymphatic filariasis	Parasite
	West Nile fever	Virus
Aquatic snails	Schistosomiasis (bilharziasis)	Parasite
Blackflies	Onchocerciasis (river blindness)	Parasite
Fleas	Plague (transmitted from rats to humans)	Bacteria
	Tungiasis	Ectoparasite
Lice	Typhus	Bacteria
	Louse-borne relapsing fever	Bacteria
Sandflies	Leishmaniasis	Parasite
	Sandfly fever (phlebotomus fever)	Virus
Ticks	Crimean-Congo haemorrhagic fever	Virus
	Lyme disease	Bacteria
	Relapsing fever (borreliosis)	Bacteria
	Rickettsial diseases (eg: spotted fever and Q fever)	Bacteria
	Tick-borne encephalitis	Virus
	Tularaemia	Bacteria
Triatome bugs	Chagas disease (American trypanosomiasis)	Parasite
Tsetse flies	Sleeping sickness (African trypanosomiasis)	Parasite

Table 1. Common vector-borne diseases (World Health Organization)[1]

1.2 Mosquito-borne diseases

Mosquito-borne diseases are the deadliest of all vector-borne diseases.[6] The reason mosquitos are so infectious is due to their need to consume a blood meal. Female mosquitos require a blood meal in order to obtain nutrients essential for laying eggs.[7] When females consume the blood meal from an infected host they carry the disease. The disease is spread to a new host throughout their life cycle every time they take a blood meal. In contrast to humans, mosquitos act as a vector to spread the disease, but are not affected by it.[8] *Aedes* mosquitos are responsible for spreading Dengue, Chikungunya, Rift Valley fever, Yellow fever, and Zika virus. *Culex* mosquitos are responsible for spreading West Nile virus and Japanese encephalitis and *Anopheles* mosquitos are responsible for spreading Malaria. *Aedes, Anopheles*, and *Culex* mosquitos are all capable of spreading the parasitic disease Lymphatic filariasis.[1,9] While Malaria is the deadliest of all mosquito-borne diseases, Dengue is the most widespread.

1.2.1 Malaria

According to the WHO, Malaria, a parasitic disease caused by *Plasmodium* is spread by *Anopheles* mosquitos and accounts for an estimated 249 million cases and 608,000 deaths in 2022.[10] Malaria disproportionately affects the WHO African Region totally 94 % of all malaria cases (233 million) and 95 % of deaths (580,000). 80 % of those deaths were

in children under 5 years old. The 4 countries that experience the most deaths from malaria are Nigeria (26.8 %), Democratic Republic of Congo (12.3 %), Uganda (5.1 %), and Mozambique (4.2 %). The 5 species of *Plasmodium* that cause malaria in humans is *P. falciparum, P. vivax, P. malariae*, and *P. ovale. P. falciparum* is the deadliest species of parasites that cause malaria and the most prevalent in Africa. Mild symptoms can progress to severe illness or death within 24 hours if left untreated. Mild symptoms include fever and flu-like symptoms such as chills, headache, muscle aches, nausea, vomiting, and diarrhea. Severe symptoms of the disease include extreme tiredness or fatigue, impaired consciousness, convulsions, difficulty breathing, dark or bloody urine, jaundice, and abnormal bleeding.[10,11]

How severe the disease progresses depends on the life cycle of the parasite in the human body. Disease is only caused after the parasite leaves the liver and invades red blood cells. Several steps before this must take place for this to occur. Once an infected mosquito bites a human host, sporozoites of the *Plasmodium* parasite are injected into the subcutaneous tissue or sometimes the blood stream of the host.[12] The sporozoites make their way to the liver and infect liver cells. Sporozoites mature into schizonts and rupturing the cells to release merozoites. The parasite then undergoes asexual replication in erythrocytes (red blood cells). The parasite can differentiate into gametocytes for sexual reproduction. Once in the blood stream the parasite can be spread to the mosquito after a blood meal.

Treatment of malaria depends on the species of *Plasmodium*. Two medicines for the treatment of malaria are derived from natural products, Quinine from the *Cinchona* plant and Artemisinin from the Qinghao plant (*Artemisia annua*).[13] Artemisinin-based combination therapy is recommended for *P. falciparum*, while Chloroquine is recommended for *P. vivax*. Primaquine can also be added to treatment to prevent relapse with *P. vivax* or *P. ovale* infections.[10] Combination-based therapies have had the most success in the treatment of malaria. Although preventing transmission of the parasite from mosquitos is the best way to prevent malaria (vector control), there are preventative medicines that can be taken. Chemoprophylaxis can be used for travelers 2-3 weeks before departure. Preventive chemotherapy is a full treatment for malaria that may be used to prevent malaria in populations that are most susceptible including those who are immunocompromised, pregnant, or school-aged children. Vaccines have been under development for decades, but showed low efficacy for preventing malaria. In 2021, the RTS,S/AS01 vaccine has shown to significantly reduce malaria, especially severe malaria in young children for *P. falciparum* infections. As of October 2023, the WHO recommends a second vaccine R21/Matrix-M. It is expected to be released across Africa as soon as possible.[10]

A major concern in the treatment of malaria is drug resistance. Arteminisin combination therapy (ACT) has become recommended due to the resistance of previously effective drugs chloroquine, sulfadoxine-

pyrimethamine, and amodiaquine. In the last 10 years, partial Artemisinin resistance has been observed in the Greater Mekong subregion as well as Eritrea, Rwanda, Uganda, and Tanzania. Mass drug administration has been effective at treating and preventing malaria, but it has also led to an increase in resistance.

1.2.2 Dengue

Although malaria is still a major global problem, Dengue has become the most widespread disease transmitted by mosquitos.[3,14] Dengue, a virus spread by *Aedes* mosquitos, put 3.9 billion people at risk in 129 countries.[1] That's a major global burden affecting approximately half of the world's population.[15] The number of infections as increased drastically in the last two decades. The WHO received reports of 505,430 cases of Dengue in 2000 and 5.2 million in 2019.[16] The WHO predicts there are 390 million viral infections of Dengue resulting in approximately 96 million symptomatic infections and 40,000 deaths annually.[1,16,17] Since Dengue is often asymptomatic and misdiagnosed as febrile disease, many cases are under-reported.[18]

1.3 Favorable climates for mosquitos

As mentioned previously for vector-borne diseases, mosquito-borne diseases are more prevalent in tropical or sub-tropical climates. Factors that

are favorable for mosquito survival are warm temperatures between 23-29°C, higher humidity, heavy rainfall, and lower altitudes.[5,19] With this being said, as temperatures rise and climate changes to more favorable conditions, mosquitos are moving to regions that were once less favorable.[4] Since female mosquitos lay their eggs in water, any standing water poses a risk for a mosquito breeding ground.[20] In addition to weather, areas that are socioeconomically disadvantaged, have a higher risk of contracting mosquito-borne diseases.[1] This includes access to health care and treatment, government infrastructure to prevent and treat mosquito-borne diseases, and proper housing and nutrition.

1.4 Vector control method for mosquitoes

Vector control is one of the most effective ways of preventing mosquito-borne diseases. Vector control refers to preventing the vector, in this case, mosquitos from infecting humans. There are 3 major types of vector control: physical, biological, and chemical. In order to understand how to control mosquitos, it is important to understand the life cycle of a mosquito. All mosquitos experience 4 major stages during their life cycle: egg, larva, pupa, and adult. *Anopheles* female mosquitos lay their eggs in water (Figure 1.1). The eggs hatch 2-3 days later into larvae. The larvae take 4-10 days to develop into pupae. After 2-3 days the pupae develop into adult mosquitos that are able to fly. The length of these steps varies slightly between each

species. Water plays a significant role in a mosquito's life cycle. Mosquitos spend 3 of their life stages developing in water before they are able to leave and fly away as adult mosquitos. This is why mosquitos prefer a humid and wet environment with heavy rainfall.

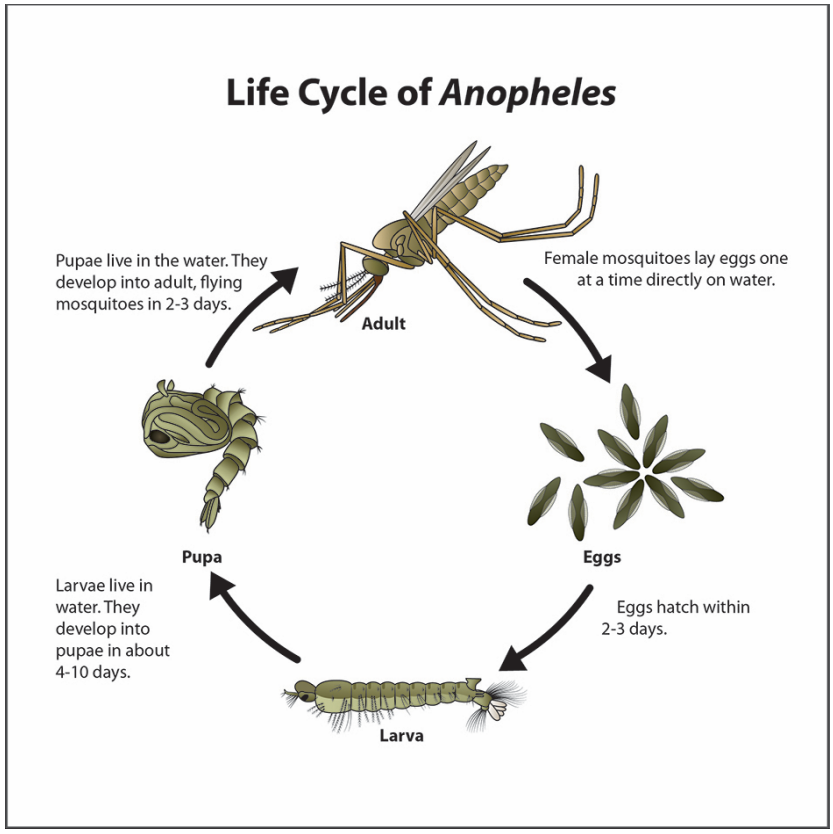


Figure 1.1. The life cycle of an *Anopheles* mosquito (CDC)[21]
Mosquitos have 4 major life stages: egg, larva, pupa, and adult. Female mosquitos lay their eggs in water. Eggs hatch 2-3 days later into larvae. The larvae take 4-10 days to develop into pupae. After 2-3 days the pupae develop into adult mosquitos that are able to fly.

1.4.1 Physical vector control

Physical vector control refers to physically blocking the mosquito from taking a blood meal. Many mosquitos tend to stay in dark places during the day and come out at night. Using mosquito nets while sleeping and having adequate housing that prevents mosquitos from entering into the building prevent biting at night. Often insecticides are sprayed on nets and buildings to further deter the mosquito from biting. Another common practice is removing standing water away from housing in order to prevent attracting female mosquitos since they need to lay their eggs in water.

1.4.2 Biological vector control

Biological larvicides include introducing bacteria that produce proteins that are lethal to the larvae or are a known predator of larvae. An example of a species used as a biological larvicide is *Bacillus thuringiensis israelensis*, a naturally occurring soil bacteria. *Bacillus thuringiensis israelensis* produces six δ-endotoxins that form a parasporal crystalline body.[22] Once inside the larva, the toxins are able to create ion channels in epithelial cells that eventually lead to cell lysis and death of the larva.[23] *Bacillus thuringiensis israelensis* has been used for decades and is non-toxic to humans. *Gambusia*, a species of fish that feed on mosquito larvae, may also be introduced into bodies of water to reduce the mosquito population.

More controversial methods include the sterilization of male mosquitos by irradiation, genetically modifying mosquitos, and infecting mosquitos with *Wolbachia*.[24] Genetic modification (GM) of mosquitos includes a self-limiting gene that prevents the GM mosquito's progeny from surviving to adulthood. *Wolbachia,* is a bacterium that can be introduced to mosquito eggs. Only the male mosquitos with *Wolbachia* are released into the environment. If a male mosquito has *Wolbachia* and mates with a female, her eggs will not hatch. Therefore, preventing new mosquitos from propagating.

1.4.3 Insecticide use for chemical vector control

Chemical vector control is the use of chemicals such as pesticides or insecticides in order to prevent or kill mosquitos from biting. This type of vector control is very effective and is used around the world. Types of chemical vector control include the spraying of pesticides outside and inside of houses (IRS) or buildings, on mosquito nets (ITNs), in bodies of water, and space spraying of areas that are largely affected by mosquitos.[25] Types of insecticide classes used for vector control include pyrethroids, carbamates, organochlorides, organophosphates, neonicotinoids, butenolides, pyrroles, spinosyns, and juvenile hormone (JH) mimics (Figure 1.2). Insecticide treated nets (ITNs) are often sprayed with pyrethroids, while pyrethroids, carbamates, organophosphates, organochlorines are used for indoor residual spraying (IRS). Treating bodies of water with larvae include the use of

organophosphates, juvenile hormone mimics, and spinosyns. Pyrethroids and organophosphates are used for space spraying.[25] As of 2017, the WHO approved the use of neonicotinoids and butenolides for IRS.

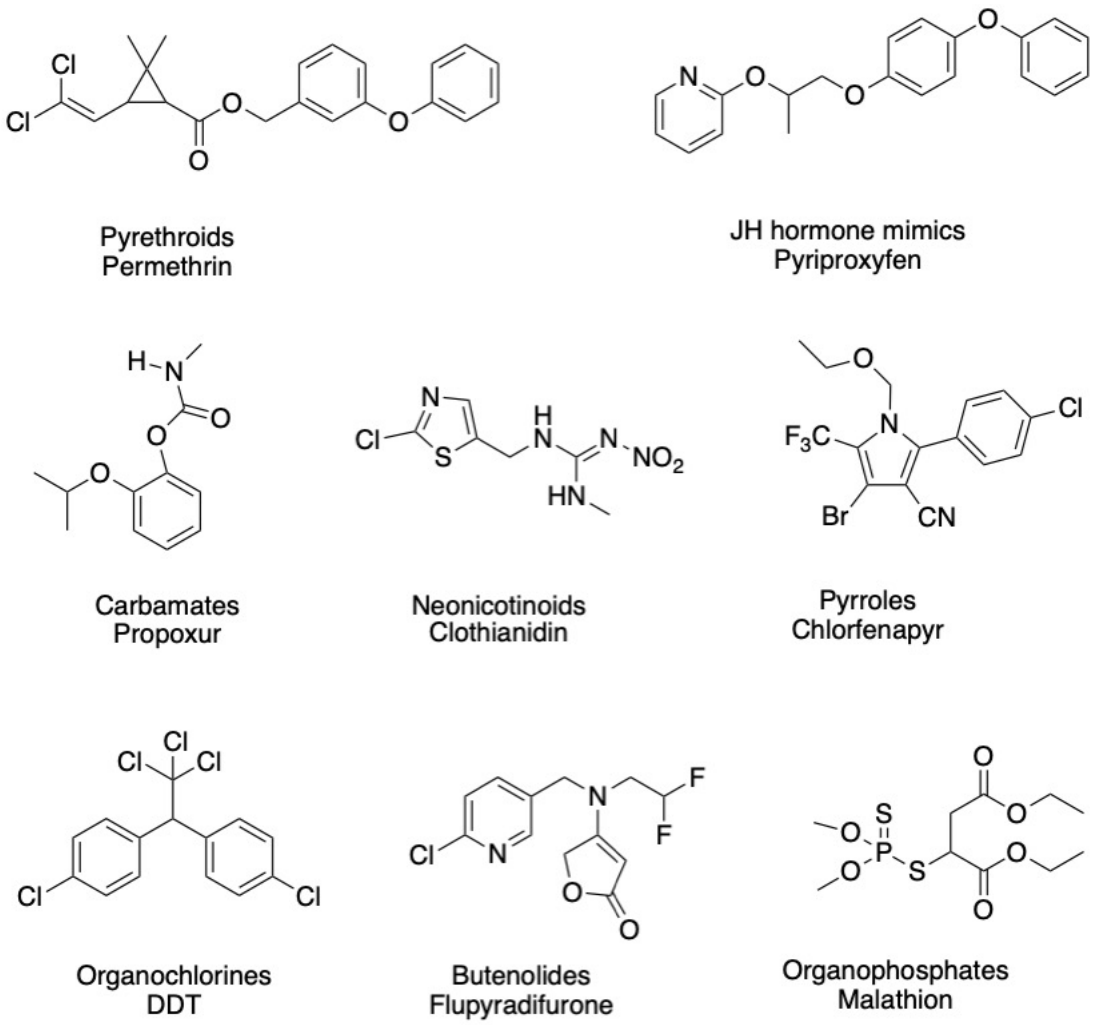


Figure 1.2 Insecticide classes and a molecular structure from each class used against mosquitos.
Insecticide classes used against mosquitos are (Top) pyrethroids, JH hormone mimics, carbamates, neonicotinoids, pyrroles, organochlorines, butenolides, and organophosphates.

Types of pyrethroids include alpha-cypermethrin, cyfluthrin, deltamethrin, etofenprox, lambda-cyhalothrin, permethrin, transfluthrin, metofluthrin, and prallethrin. Carbamates include bendiocarb, carbosulfan, and propoxur. Organochlorines include DDT and dieldrin, and

organophosphates include fenitrothion, chlorpyrifos-ethyl, malathion, and pirimiphos-methyl. Other insecticides include clothianidin from neonicotinoids, flupyradifurone from butenolides, chlorfenapyr from pyrroles, and pyriproxyfen from JH hormone mimics.[26] Most of these insecticides act as neurotoxins. Their targets include voltage-dependent sodium channels, acetylcholineeterases, and receptors of γ-aminobutyric acid.[27] Two major insecticides classes of pyrethroids and organochlorines (DDT) target voltage-dependent sodium gated channels.[28]

Insecticides can be further categorized into two classes, adulticides and larvicides. Larvicides are used to prevent adult mosquitos from developing into adults and can be both biological and chemical. In contrast to adulticides which are sprayed in the air, larvicides are added to standing water. A chemical larvicide such a Methoprene, is an insect-growth regulator. The larvae are unable to mature into adult mosquitos in its presence.

1.5 Mosquito resistance of insecticides

As mentioned previously with antimalarials, resistance is a major global concern. The World Health Organization has worked extensively to track resistance to insecticides and set up recommendations on how to monitor and intervene.[25] Resistance to the four oldest and most widely used insecticides pyrethroids, carbamates, organophosphates, and organochlorides have been observed globally.[29,30] It is unknown how

widespread resistance is for the newer classes of insecticides. Since many of these insecticides are also used for agriculture, it is likely resistance will develop overtime. According to the WHO “resistance is caused by the evolutionary selection of specific genetic mutations in mosquitos that allow them to survive exposure to insecticides”.[25] This could include mutations in voltage-sensitive sodium channels (kdr) that are targeted by some insecticides.[31–33] Other mechanisms could include upregulating production of enzymes such as cytochrome P450, monooxygenases, esterases, carboxylesterases, and glutathione S-transferases that are involved detoxification.[34–36] Other studies have found thickening of the cuticle is associated with resistance due to the reduction of absorption of insecticides.[37–39] Because of resistance, it is important to find new classes of insecticides with different targets specific to mosquitos.

1.6 Detrimental effects of insecticides

Insecticides are a very effective vector control method, especially in combination with physical vector control methods. Unfortunately, insecticides have had a long history of detrimental effects on other insects, animals, and humans. For example, DDT which was once widely used is now banned for all uses except for malaria control. It is still on the WHO list of insecticides recommended for IRS to prevent malaria. It states the benefit of preventing malaria outweighs the health and environmental risk.[40] DDT is often used for

the control of pyrethroid resistant mosquitos.[41] Rachel Carson, a environmental scientist, published the book Silent Spring in 1962 outlining the detrimental effects of DDT to wildlife based on scientific evidence.[42] Her research and testimony at two congressional hearings led to the development of the EPA in 1970 and several laws to protect the environment and human health.[43] Ten years after her publication, the EPA banned the use of DDT in 1972.[40] High exposure to DDT can cause vomiting, tremors, and seizures. Studies in mice have shown liver and reproductive harm. It is also a suspected carcinogen. Long term effects have shown mutations caused by DDT have been passed down through generations.[44]

1.7 Summary

Mosquitos are the deadliest animal in the world. Mosquitos carry multiple deadly diseases such as malaria and dengue that disproportionately affect the poorest communities and children under 5. Mosquitos in particular are high transmitting vectors due to their population density and the nature of their lifecycle, requiring a blood meal. Currently we use a combination of physical, biological, and chemical control methods to limit the spread of mosquito-borne diseases. All of these control methods have limitations. Physical control methods such as mosquito nets or adequate housing that provides a secure barrier from the outside environment are scarce in poorer communities and lead to higher disease rates. Biological control methods are

controversial and it is unknown what long-term effects these control methods could have on the ecosystem. On the other hand, chemical control methods like insecticides are effective but they are often non-specific and kill many other organisms. These off-target effects are a significant problem that contributes to the development of resistance to insecticides. Insecticides have been shown to be bad for the environment by causing damage to ecosystems and directly harming humans by promoting both acute and chronic health issues.

Some of the strategies to limit the development of resistance include combination therapy where multiple drugs are used to target alternate pathological mechanisms. Natural products also offer a promising alternative strategy because they are naturally derived from organisms in the environment that produce them as a defense mechanism to a specific environmental challenge. This allows humans to harness the evolutionarily derived function and specificity of natural products and leverage them toward the challenges that we face, without the risk of introducing man-made chemicals into the environment. This dissertation is aimed at finding a naturally derived product that selectively kills mosquitos for the development of a mosquito-specific insecticide for the prevention of mosquito-borne diseases.

1.8 References

(1) *Vector-borne diseases*. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases (accessed 2023-11-30).

(2) *Diseases Carried by Vectors | CDC*. https://www.cdc.gov/climateandhealth/effects/vectors.htm (accessed 2023-11-30).

(3) Obradovic, Z.; Smjecanin, E.; Pindzo, E.; Omerovic, H.; Cibo, N. A LITERATURE REVIEW ON VECTOR BORNE DISEASES. *Int J Med Rev Case Rep* **2022**, *6* (Reports in Surgery, Orthopaedi), 1. https://doi.org/10.5455/IJMRCR.172-1639404085.

(4) Caminade, C.; McIntyre, K. M.; Jones, A. E. Impact of Recent and Future Climate Change on Vector-Borne Diseases. *Annals of the New York Academy of Sciences* **2019**, *1436* (1), 157–173. https://doi.org/10.1111/nyas.13950.

(5) Chala, B.; Hamde, F. Emerging and Re-Emerging Vector-Borne Infectious Diseases and the Challenges for Control: A Review. *Frontiers in Public Health* **2021**, *9*.

(6) Abdullah Shaukat, M.; Ali, S.; Saddiq, B.; Waqar Hassan, M.; Ahmad, A.; Kamran, M. Effective Mechanisms to Control Mosquito Borne Diseases: A Review. **2019**, *4*, 21–30.

(7) Nazareth, T.; Seixas, G.; Sousa, C. A. Climate Change and Mosquito-Borne Diseases. In *Climate Change and Health: Improving Resilience and Reducing Risks*; Leal Filho, W., Azeiteiro, U. M., Alves, F., Eds.; Climate Change Management; Springer International Publishing: Cham, 2016; pp 187–199. https://doi.org/10.1007/978-3-319-24660-4_12.

(8) Prevention, C.-C. for D. C. and. *CDC - Malaria - About Malaria - Biology*. https://www.cdc.gov/malaria/about/biology/index.html (accessed 2023-12-01).

(9) CDC. *Mosquitoes and Diseases: A-Z | CDC*. Centers for Disease Control and Prevention. https://www.cdc.gov/mosquitoes/about/diseases.html (accessed 2023-11-30).

(10) *Fact sheet about malaria*. https://www.who.int/news-room/fact-sheets/detail/malaria (accessed 2023-12-04).

(11) Prevention, C.-C. for D. C. and. *CDC - Malaria - About Malaria*. https://www.cdc.gov/malaria/about/index.html (accessed 2023-12-04).

(12) Miller, L. H.; Baruch, D. I.; Marsh, K.; Doumbo, O. K. The Pathogenic Basis of Malaria. *Nature* **2002**, *415* (6872), 673–679. https://doi.org/10.1038/415673a.

(13) Talapko, J.; Škrlec, I.; Alebić, T.; Jukić, M.; Včev, A. Malaria: The Past and the Present. *Microorganisms* **2019**, *7* (6), 179. https://doi.org/10.3390/microorganisms7060179.

(14) Benelli, G.; Mehlhorn, H. Declining Malaria, Rising of Dengue and Zika Virus: Insights for Mosquito Vector Control. *Parasitol Res* **2016**, *115* (5), 1747–1754. https://doi.org/10.1007/s00436-016-4971-z.

(15) *Population Clock*. https://www.census.gov/popclock/ (accessed 2023-11-30).

(16) *Dengue and severe dengue*. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue (accessed 2023-12-04).

(17) Brady, O. J.; Gething, P. W.; Bhatt, S.; Messina, J. P.; Brownstein, J. S.; Hoen, A. G.; Moyes, C. L.; Farlow, A. W.; Scott, T. W.; Hay, S. I. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLOS Neglected Tropical Diseases* **2012**, *6* (8), e1760. https://doi.org/10.1371/journal.pntd.0001760.

(18) Bhatt, S.; Gething, P. W.; Brady, O. J.; Messina, J. P.; Farlow, A. W.; Moyes, C. L.; Drake, J. M.; Brownstein, J. S.; Hoen, A. G.; Sankoh, O.; Myers, M. F.; George, D. B.; Jaenisch, T.; Wint, G. R. W.; Simmons, C. P.; Scott, T. W.; Farrar, J. J.; Hay, S. I. The Global Distribution and Burden of Dengue. *Nature* **2013**, *496* (7446), 504–507. https://doi.org/10.1038/nature12060.

(19) Mordecai, E. A.; Caldwell, J. M.; Grossman, M. K.; Lippi, C. A.; Johnson, L. R.; Neira, M.; Rohr, J. R.; Ryan, S. J.; Savage, V.; Shocket, M. S.; Sippy, R.; Stewart Ibarra, A. M.; Thomas, M. B.; Villena, O. Thermal

Biology of Mosquito-Borne Disease. *Ecology Letters* **2019**, *22* (10), 1690–1708. https://doi.org/10.1111/ele.13335.

(20) US EPA, O. *General Information about Mosquitoes*. https://www.epa.gov/mosquitocontrol/general-information-about-mosquitoes (accessed 2023-11-30).

(21) *Life Cycle of Anopheles Species Mosquitoes | Mosquitoes | CDC*. https://www.cdc.gov/mosquitoes/about/life-cycles/anopheles.html (accessed 2023-12-05).

(22) Ben-Dov, E. Bacillus Thuringiensis Subsp. Israelensis and Its Dipteran-Specific Toxins. *Toxins* **2014**, *6* (4), 1222–1243. https://doi.org/10.3390/toxins6041222.

(23) Donovan, W. P.; Rupar, M. J.; Slaney, A. C.; Malvar, T.; Gawron-Burke, M. C.; Johnson, T. B. Characterization of Two Genes Encoding Bacillus Thuringiensis Insecticidal Crystal Proteins Toxic to Coleoptera Species. *Applied and Environmental Microbiology* **1992**, *58* (12), 3921–3927. https://doi.org/10.1128/aem.58.12.3921-3927.1992.

(24) CDC. *Genetically Modified Mosquitoes | CDC*. Centers for Disease Control and Prevention. https://www.cdc.gov/mosquitoes/mosquito-control/community/emerging-methods/genetically-modified-mosquitoes.html (accessed 2023-11-28).

(25) Manual for Monitoring Insecticide Resistance in Mosquito Vectors and Selecting Appropriate Interventions.

(26) *Malaria vector control*. https://www.who.int/teams/global-malaria-programme/prevention/vector-control/insecticide-resistance (accessed 2023-12-06).

(27) Sokhna, C.; Ndiath, M. O.; Rogier, C. The Changes in Mosquito Vector Behaviour and the Emerging Resistance to Insecticides Will Challenge the Decline of Malaria. *Clinical Microbiology and Infection* **2013**, *19* (10), 902–907. https://doi.org/10.1111/1469-0691.12314.

(28) Karunaratne, S. H. P. P.; Silva, W. A. P. P. D.; Weeraratne, T. C.; Surendran, S. N. Insecticide Resistance in Mosquitoes: Development, Mechanisms and Monitoring. **2018**, *47* (4), 299. https://doi.org/10.4038/cjs.v47i4.7547.

(29) *World malaria report 2023*. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023 (accessed 2023-12-06).

(30) *Malaria Threat Map*. https://apps.who.int/malaria/maps/threats/#/maps?theme=prevention&mapType=prevention%3A0&bounds=%5B%5B-171.06858566430066%2C-78.03037980141767%5D%2C%5B219.76100902172925%2C37.11787370810495%5D%5D&insecticideClass=PYRETHROIDS&insecticideTypes=&assayTypes=MOLECULAR_ASSAY%2CBIOCHEMICAL_ASSAY%2CSYNERGIST-INSECTICIDE_BIOASSAY&synergistTypes=&species=&vectorSpecies=

&surveyTypes=&plasmodiumSpecies=P._FALCIPARUM&drugs=DRUG_AL&mmType=1&excludeLowerPatients=false&excludeLowerSamples=false&endemicity=false&storyMode=false&storyModeStep=0&years=1978%2C2023 (accessed 2023-12-06).

(31) Li, C.-X.; Kaufman, P. E.; Xue, R.-D.; Zhao, M.-H.; Wang, G.; Yan, T.; Guo, X.-X.; Zhang, Y.-M.; Dong, Y.-D.; Xing, D.; Zhang, H.-D.; Zhao, T.-Y. Relationship between Insecticide Resistance and Kdr Mutations in the Dengue Vector Aedes Aegypti in Southern China. *Parasites & Vectors* **2015**, *8* (1), 325. https://doi.org/10.1186/s13071-015-0933-z.

(32) Matambo, T. S.; Abdalla, H.; Brooke, B. D.; Koekemoer, L. L.; Mnzava, A.; Hunt, R. H.; Coetzee, M. Insecticide Resistance in the Malarial Mosquito Anopheles Arabiensis and Association with the Kdr Mutation. *Medical and Veterinary Entomology* **2007**, *21* (1), 97–102. https://doi.org/10.1111/j.1365-2915.2007.00671.x.

(33) Reimer, L.; Fondjo, E.; Patchoké, S.; Diallo, B.; Lee, Y.; Ng, A.; Ndjemai, H. M.; Atangana, J.; Traore, S. F.; Lanzaro, G.; Cornel, A. J. Relationship Between Kdr Mutation and Resistance to Pyrethroid and DDT Insecticides in Natural Populations of Anopheles Gambiae. *Journal of Medical Entomology* **2008**, *45* (2), 260–266. https://doi.org/10.1093/jmedent/45.2.260.

(34) Yang, T.; Liu, N. Genome Analysis of Cytochrome P450s and Their Expression Profiles in Insecticide Resistant Mosquitoes, Culex

Quinquefasciatus. *PLOS ONE* **2011**, *6* (12), e29418. https://doi.org/10.1371/journal.pone.0029418.

(35) Denholm, I.; Pickett, J. A.; Devonshire, A. L.; Hemingway, J.; Hawkes, N.; Prapanthadara, L.; Jayawardenal, K. G. I.; Ranson, H. The Role of Gene Splicing, Gene Amplification and Regulation in Mosquito Insecticide Resistance. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **1998**, *353* (1376), 1695–1699. https://doi.org/10.1098/rstb.1998.0320.

(36) Hemingway, J. The Molecular Basis of Two Contrasting Metabolic Mechanisms of Insecticide Resistance. *Insect Biochemistry and Molecular Biology* **2000**, *30* (11), 1009–1015. https://doi.org/10.1016/S0965-1748(00)00079-5.

(37) Wood, O. R.; Hanrahan, S.; Coetzee, M.; Koekemoer, L. L.; Brooke, B. D. Cuticle Thickening Associated with Pyrethroid Resistance in the Major Malaria Vector Anopheles Funestus. *Parasites & Vectors* **2010**, *3* (1), 67. https://doi.org/10.1186/1756-3305-3-67.

(38) Yahouédo, G. A.; Chandre, F.; Rossignol, M.; Ginibre, C.; Balabanidou, V.; Mendez, N. G. A.; Pigeon, O.; Vontas, J.; Cornelie, S. Contributions of Cuticle Permeability and Enzyme Detoxification to Pyrethroid Resistance in the Major Malaria Vector Anopheles Gambiae. *Sci Rep* **2017**, *7* (1), 11091. https://doi.org/10.1038/s41598-017-11357-z.

(39) Samal, R. R.; Kumar, S. Cuticular Thickening Associated with Insecticide Resistance in Dengue Vector, Aedes Aegypti L. *Int J Trop Insect Sci* **2021**, *41* (1), 809–820. https://doi.org/10.1007/s42690-020-00271-z.

(40) US EPA, O. *DDT - A Brief History and Status*. https://www.epa.gov/ingredients-used-pesticide-products/ddt-brief-history-and-status (accessed 2023-12-06).

(41) van den Berg, H.; Manuweera, G.; Konradsen, F. Global Trends in the Production and Use of DDT for Control of Malaria and Other Vector-Borne Diseases. *Malaria Journal* **2017**, *16* (1), 401. https://doi.org/10.1186/s12936-017-2050-2.

(42) Carson, R. *Silent Spring*; Houghton Mifflin: Boston, 1962.

(43) *Legacy of Rachel Carsons Silent Spring National Historic Chemical Landmark*. American Chemical Society. https://www.acs.org/education/whatischemistry/landmarks/rachel-carson-silent-spring.html (accessed 2023-12-06).

(44) Kabasenche, W. P.; Skinner, M. K. DDT, Epigenetic Harm, and Transgenerational Environmental Justice. *Environmental Health* **2014**, *13* (1), 62. https://doi.org/10.1186/1476-069X-13-62.

Chapter 2

Discovery of a mosquito-specific toxin NP-34

2.1 Introduction

2.1.1 Natural products in drug discovery

Natural products have a played a major role in drug discovery.[1] Natural products, also known as secondary metabolites, are compounds produced by plants, animals, or microorganisms.[2] Unlike primary metabolites, secondary metabolites are not required for survival.[2] It is hypothesized these secondary metabolites are produced as an evolutionary advantage to the organism.[3] For example, the bacteria *Streptomyces* produces the most antibiotics of any genus known so far.[4] These antibiotics are thought to protect itself by killing other organisms.[5] In addition to antibiotics, natural products or derivatives of natural products have been used to treat a wide variety of diseases and ailments such as malaria, cancer, viral infections, high cholesterol, pain, inflammation, etc.[1]

Along with treating disease, natural products have also been used effectively as insecticides. Spinosad consisting of spinosyn A and B produced by *Saccharopolyspora spinosa*, pyrethrum consisting of pyrethrin I and II, cinerin I and II, and jasmolin I and II from *Tanacetum cinerariaefolium*, rotenone from *Lonchocarpus utilis* and *Lonchocarpus urucu*, cevadine from *Schoenocaulon officinale,* avermectin consisting of avermectin B1a and B1b from *Streptomyces avermitilis*, milbemycin consisting of milbemycin A3 and A4 from *Streptomyces hygroscopicus*, and azadirachtin A from *Azadirachta*

indica are all natural products used as insecticides (Figure 2.1).[6] Many of these natural products have been approved as an organic insecticide for because they are naturally derived. Although many of these are considered safe for humans, they are nonselective against insects. For example, pyrethrum is a neurotoxin that kills over 40 insects.[6] With this in mind, the goal of this project is to create a safe and effective insecticide only toxic to mosquitos.

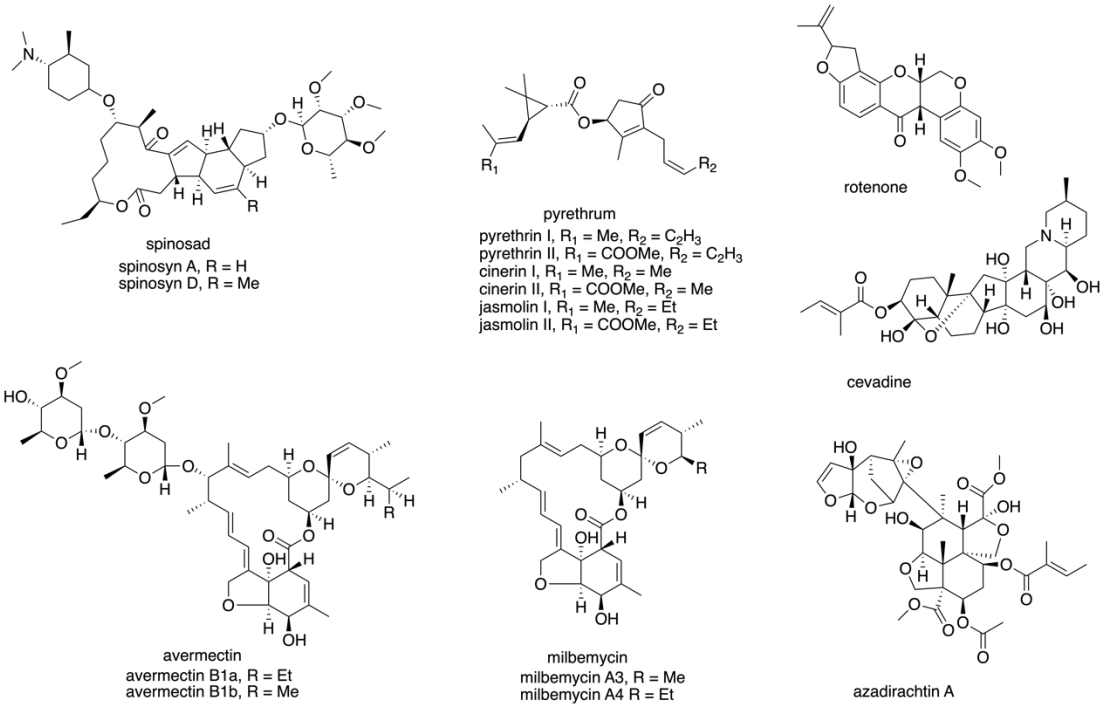


Figure 2.1. Natural products used as insecticides
(Top) Spinosad consisting of spinosyn A and B produced by *Saccharopolyspora spinosa*, pyrethrum consisting of pyrethrin I and II, cinerin I and II, and jasmolin I and II from *Tanacetum cinerariaefolium*, rotenone from *Lonchocarpus utilis* and *Lonchocarpus urucu*, cevadine from *Schoenocaulon officinale,* avermectin consititng of avermectin B1a and B1b from *Streptomyces avermitilis*, milbemycin consisting of milbemycin A3 and A4 from *Streptomyces hygroscopicus*, and azadirachtin A from *Azadirachta indica*.

2.1.2 High-throughput screening to determine a mosquito-specific toxin

To determine a toxin that is specific to mosquitoes, a high-throughput screen designed by the Abrams lab was performed at UT Southwestern.[7] The initial assay included cell lines from mosquitoes (*Culex*), *Drosophila*, and *Spodoptera*. These cell lines were chosen to determine a toxin specific to the mosquito cells while remaining nontoxic to the *Spodoptera* and *Drosophila* cell lines. Our fraction library consisting of over 6,000 natural product fractions

were screened against these cell lines. The toxicity was determined by percent mortality through CellTiter-Glo. CellTiter-Glo is a luminescent cell viability assay that determines viability through the detection of ATP. A natural product fraction from SNC-034 or *Streptomyces malachitospinus* showed potent activity while maintaining selectivity to mosquitoes. The collection and isolation of this strain is explained further in the methods section. This fraction was purified further to a single compound NP-34. At 50 nM, purified NP-34 achieved 80 % mortality against mosquito cells vs. less than 10 % against *Drosophila* and *Spodoptera* cells (Figure 2.2A). Through x-ray crystallography, circular dichroism, and NMR, the compound was determined to be a boronated polyketide macrolide from the aplasmomycin family (Figure 2.2B).

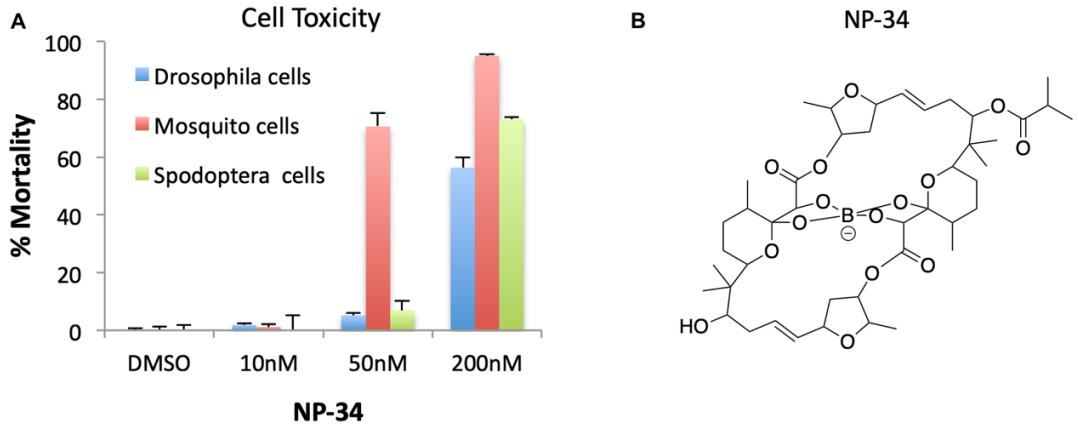


Figure 2.2. Natural product NP-34 maintains toxicity and selectivity towards mosquito cells.
A. Toxicity of NP-34 at 10 nM, 50 nM, and 200 nM in Drosophila, Mosquito (Culex), and Spodoptera cells. At 50 nM, NP-34 achieves approximately 80 % mortality of mosquito cells, while only killing less than 10 % of Spodoptera and mosquito cells. B. Chemical structure of NP-34, a boronated polyketide macrolide from the aplasmomycin family

The safety of NP-34 was further tested in human cell lines of epithelial and neuroendocrine origin (Figure 2.3). All cells had a greater than 90 % survival rate at 20 uM NP-34, a concentration much higher than intial toxicity studies in mosquito cells at 50 nM (Figure 2.2). These results indicate NP-34 is safe around humans.

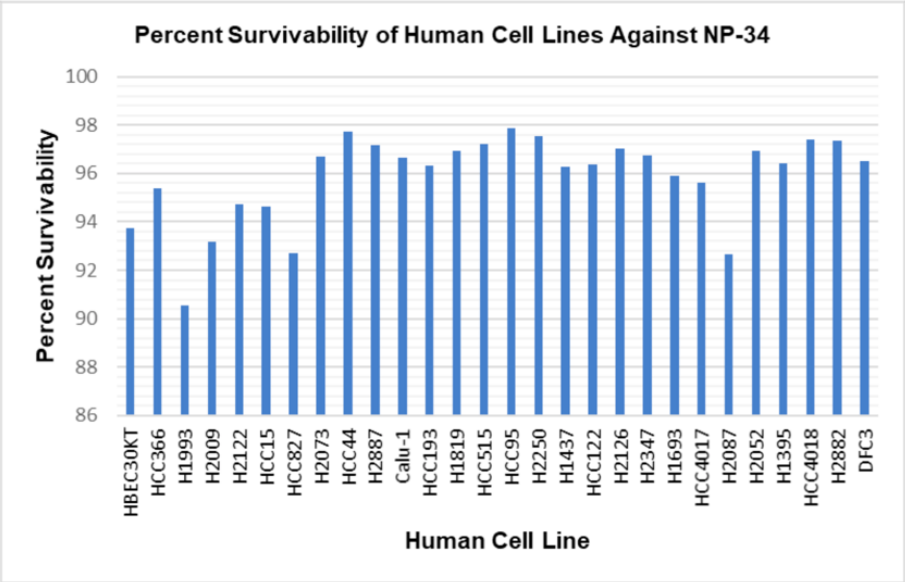


Figure 2.3. Safety of NP-34 against human epithelial and neuroendocrine cell lines.
Against 28 human cell lines, all had greater than 90 % survivability against 20 uM NP-34 normalized to a DMSO control.

2.1.3 Characterization of NP-34

Characterization of NP-34 was originally done by a previous member of the lab, Peng Fu, through ^{1}H and ^{13}C NMR. Based on a ^{1}H-^{13}C HSQC spectra, the molecule is made up of 10 methyl, 8 methylene, 4 olefinic methine, 15 aliphatic methine groups, and 1 exchangeable proton. The carbon spectrum revealed 7 additional carbon atoms comprising of 4 oxygenated sp3 hybridized quaternary carbons and 3 carbonyls. ^{1}H-^{13}C HMBC resembled a symmetrical 34-membered core like aplasmomycin. TOSCY verified two six-membered rings C-3 to C-7 and C-3' to C-7' and two tetrahydrofuran rings from C-13 to C-16 and C-13' to C-16'. Based on ^{1}H-^{13}C

HMBC and TOSCY, NP-34 is symmetrical until the side chain off C-9 and C-9', which was identified to be an isobutyrate (^{1}H 2.54 ppm to C-1", C-3", and C-4" and from C-9 to C-1". When 1H 2.54 ppm was irradiated by TOSCY, one doublet was revealed at 1.16 indicating two methyl groups at C-3" and C-4". C-9' differed from C-9 indicating an alcohol instead of an isobutyrate. This information was based on the proton at ^{1}H 5.51 ppm with ^{1}H-^{13}C HMBC correlations to C-8', C-9', and C-10', but not a ^{1}H-^{13}C HSQC correlation between 1H 5.51 ppm and C-9'.

^{1}H-^{13}C HMBC confirmed the quaternary carbons C-3 and C-3' off the boron core with ^{1}H 3.73 and 4.49 correlating to C-3 and ^{1}H 3.79 and 4.55 ppm correlating to C-3'. Although no signal was found indicating a boron through ^{1}H and ^{13}C signals, the core of the molecule, like aplasmomycin, suggests boron could be present (Figure 2.4). Further studies by Jocelyn Macho confirmed the presence of a boron through ^{11}B-NMR. She identified the presence of boron in a BO_4^- configuration at 10.5 ppm (Figure 2.10).

2.1.4 Known boronated natural products

Although not common, there are a few known boronated polyketide macrolide natural products. These macrolides are all thought to be produced by polyketide synthases (PKS) and coordinate a boron in the center of 4 oxygens (Figure 2.4). The first boronated natural product discovered was boromycin isolated from *Streptomyces antibioticus* in 1967.[8] The tartrolons,

the only annotated BGC, is produced by *Teredinibacter turnerae* discovered in the cecum of shipworms.[9] Aplasmomycin, which is the most structurally similar molecule to NP-34, was originally discovered to be produced by *Streptomyces griseus.*[10] Lastly, borophycin was discovered in an extract of blue-green algae (cyanobacterium) *Nostoc linckia.*[11]

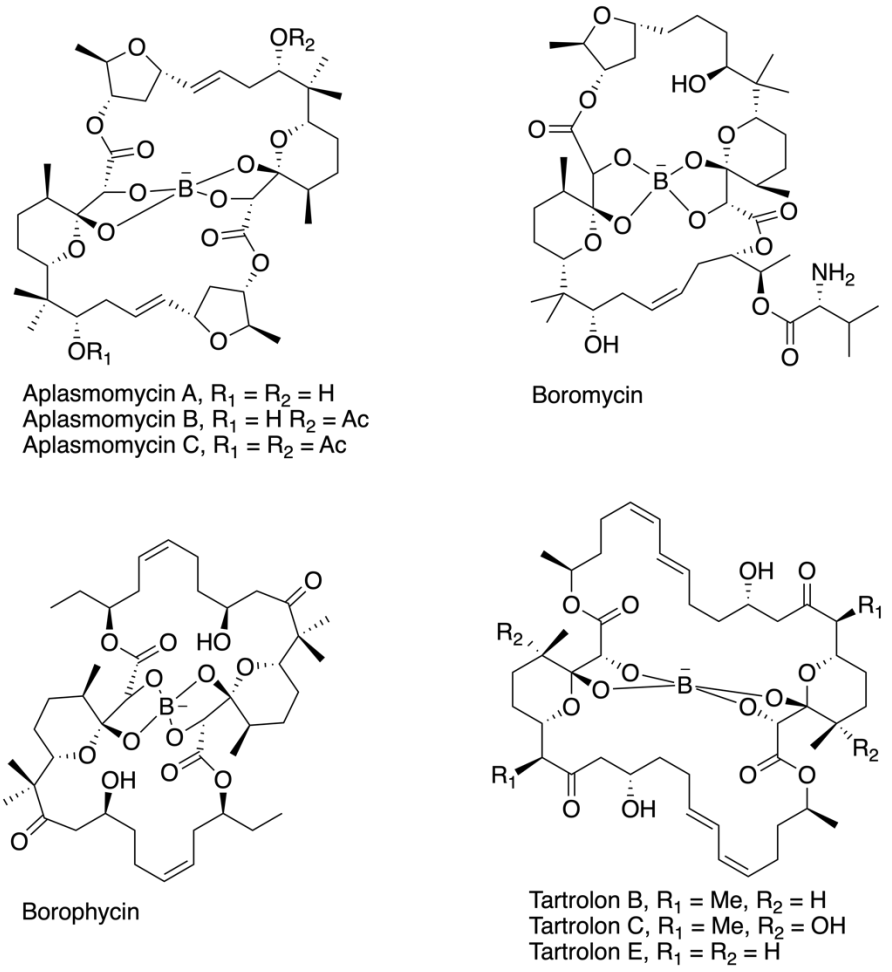


Figure 2.4. Boronated polyketide macrolide natural products. Structure of aplasmomycin A, B, and C from *Streptomyces griseus*, boromycin from *Streptomyces antibioticus*, borophycin from *Nostoc linckia*, and tartrolon B, C, and E from *Teredinibacter turnerae*.

Since the discovery of these boronated macrolides, there have been numerous studies exploring the activity of these molecules against different diseases. Although the main goal of this study was to find a mosquito-specific toxin, studies with aplasmomycin have shown antiplasmodial activity.[12] Plasmodium, as mentioned in Chapter 1, is the parasite that causes malaria spread by mosquitos. Interestingly, aplasmomycin or derivatives of aplasmomycin such as NP-34 may be toxic to both mosquitoes and plasmodium. To determine if the acyl-group in NP-34 changes the activity against mosquitoes, our collaborators tested the unacylated analog (NP34-PFb) and doubly acylated form (NP34-PFb) of NP-34 (Figure 2.5). NP34-PFb exhibited less than 10 % cell death against *Ae. aegypti* compared to 100 % cell death with NP-34 at 10 uM. NP34-PFa exhibited approximately 40 % cell death compared to approximately 90 % cell death with NP-34 at 200 nM. These results suggest the acyl group is required for activity, but only one acyl group is required. Surprisingly, the doubly acylated version (NP34-PFa) has almost half the activity as the singly acylated form of NP-34.

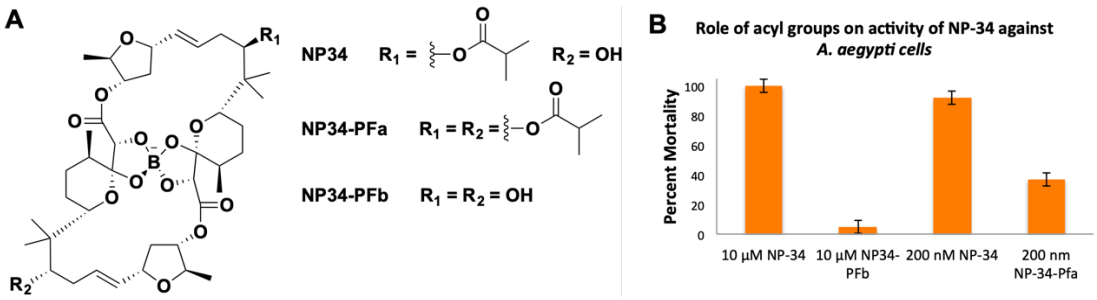


Figure 2.5. Acyl group in NP-34 is required for activity against mosquito cells
A. Molecular structure of NP-34 and doubly acylated analog PFa and unacylated analog PFb. **B**. Percent mortality of Aedes aegypti cells against NP-34, PFa, and PFb. PFb exhibited less than 10 % cell death against Ae. aegypti compared to 100 % cell death with NP-34 at 10 uM. PFa exhibited approximately 40 % cell death compared to approximately 90 % cell death with NP-34 at 200 nM.

2.1.5 NP-34 toxicity in larva and adult mosquitos

To further investigate the effectiveness of NP-34 against live mosquitos, larvae and adult mosquitos were tested. Compared to permethrin, NP-34 maintained selective toxicity to adult mosquitos *Aedes aegypti*, *Anopholes quadrimaculatus*, and *Culex quinq* over house flies (*Musca domestica*) (Figure 2.6).

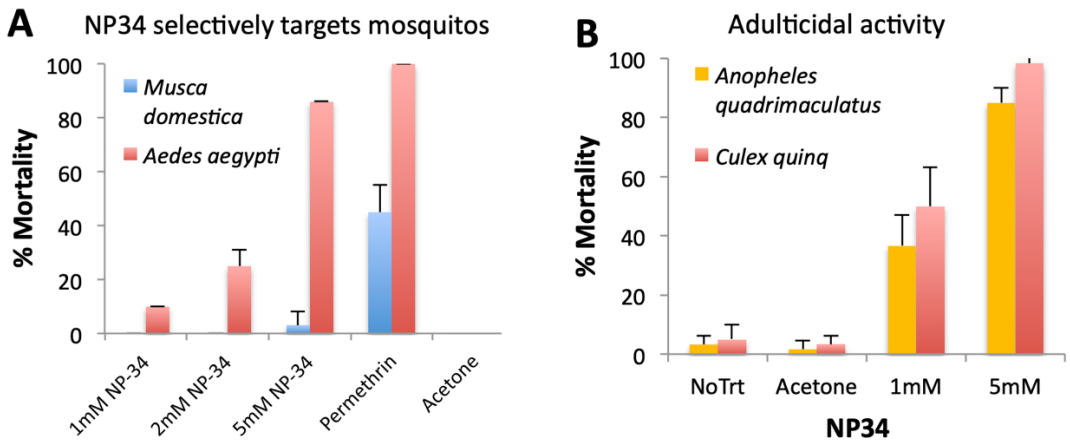


Figure 2.6. Toxicity and selectivity of NP-34 in adult mosquitos. Both studies were performed with adult mosquitoes. 0.5 uL of indicated concentration was applied topically to the thorax of the mosquito. (**A**) 1 mM, 2 mM, and 5 mM of NP-34, 26 mM permethrin (positive control), and acetone (negative control) were tested against mosquitos (*Aedes aegypti*) and house flies (*Musca domestica*). NP-34 maintained toxicity and selectivity in adult mosquitoes while permethrin had significant toxicity in house flies and mosquitos. (**B**) Two additional mosquito species were tested, *Anopheles quadrimaculatus* and *Culex quinq*. NP-34 maintained toxicity in all three mosquito strains known to carry mosquito-borne diseases.

Because some insecticides may only be active in adult mosquitos, larvae were also tested. NP-34 maintained activity in *Aedes aegypti* and *Culex quinq* larvae (Figure 2.7). Since insecticide resistance is a major concern in the prevention of mosquito-borne diseases, toxicity of NP-34 in permethrin resistant mosquitos was evaluated. NP-34 maintained toxicity in adult mosquitos and larval strains (*Aedes aegypti* R) that were resistant to permethrin (Figure 2.7).

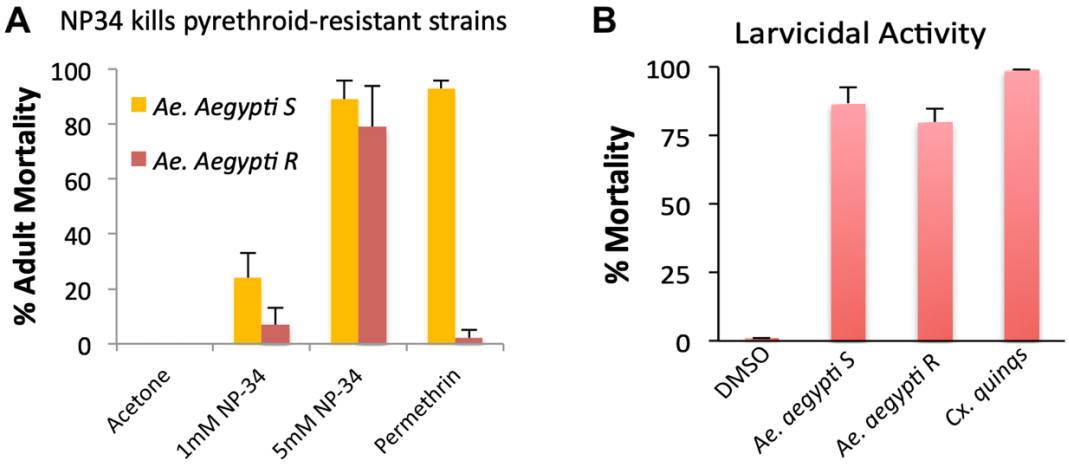


Figure 2.7. NP-34 maintains toxicity in pyrethroid-resistant adult mosquitos and larvae.
(**A**) 0.5 uL of NP-34 (1 mM and 5 mM) and 26 mM permethrin was applied topically to the thorax of adult mosquitos with a without pyrethroid resistance (*Aedes aegypti* R and S) in three trials, 20 adults each. NP-34 maintained toxicity around 80 % or greater in both strains at 5 mM, while pyrethroid-resistant mosquitos were unaffected by permethrin. (**B**) 85 uM of NP-34 was challenged against 1st intar larva of Aedes aegypti S and R and Culex cuinq in three trials, 10 animals each. Greater than 75 % mortality was observed in every strain tested.

2.1.6 In vivo acute toxicity testing of NP-34 indicated no effect on mice

To further test the safety of NP-34, preliminary toxicity studies were performed in mice. Female mice, 6 weeks old, were given 0.25 mL of NP-34 from 5.5-500 mg/kg in concentration and monitored for 30 days (Figure 2.8). The weight of the mice given NP-34 vs. the vehicle control indicated no significant difference. These results indicate NP-34 is nontoxic in mice at these concentrations. Further animal studies would be beneficial for confirming these results.

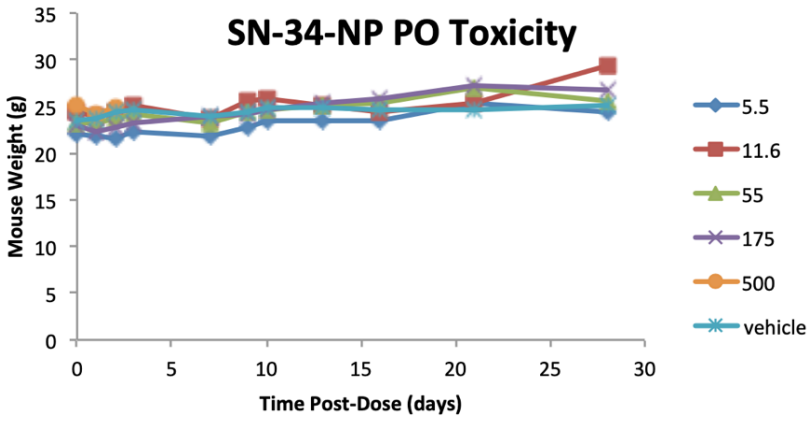


Figure 2.8. NP-34 shows no toxic effects in mice.
0.25 mL of NP-34 at concentrations of 5.5, 11.6, 55, 175, or 500 mg/kg were given orally to female mice at 6 weeks old and monitored for 30 days. The weight of the mice vs. the vehicle control indicated now effect of NP-34 in mice.

Overall, these studies indicate the effectiveness of NP-34 as a possible selective insecticide against mosquitos. The biggest experimental challenge is the production of NP-34. Since NP-34 is produced at low yields, a major goal of this project was to increase production biosynthetically. Because of the complexity of the structure, the complete synthesis of this molecule would be tedious, if at all possible. The next step was to determine how this molecule was made biosynthetically and increase production heterologously in a different host.

2.1.7 Polyketide synthases

Polyketide synthases (PKS) are observed in a wide variety of species. PKSs are evolutionarily related to fatty acid synthases of primary metabolism.[13] There are three major types of PKSs: type I, II, and III.[14] Type I

PKSs are large enzymes made of different modules. Each module performs one elongation step for the growing polyketide chain. The chain is transferred to the next module and the next unit is added to the growing polyketide chain.[15] Type I PKSs are noniterative, so each module is only visited once per polyketide made. Natural products produced by type I PKSs include erythromycin from the well studies DEBS pathway and tartrolons, a boronated polyketide macrolide with similar structure to NP-34.[9,16] In contrast, type II PKSs are iterative.[14,17] The same enzymes are used more than once for each polyketide molecule produced. Molecules produced by type II PKSs include actinorhodin, tetracyclines, anthracyclines, and aureolic acid antibiotics.[16,17] Type III PKSs, also known as the chalcone synthase-like PKSs, are homodimers of keto synthases.[18] These PKSs are also iterative, but do not contain an acyl carrier protein.[14,19] Molecules produced from these enzymes include alkylquinone, germicidin, and flaviolin.[19]

Since type I PKSs are of most interest to this study, their mechanism will be described in more detail. For type I PKSs, each module contains a ketosynthase (KS), acyl carrier protein (ACP), and an acyl transferase (AT).[15] Note that Trans-AT PKSs contain one AT enzyme that moves between modules.[13] Although not required, modules can contain tailoring enzymes such as a keto reductase (KR), dehydratase (DH), enol reductase (ER), and methyl transferase (MT). The first step of a PKS is the loading of a starter unit onto the loading module. The AT transfers an acyl group from an acyl-

Coenzyme A (acyl-CoA) onto the phosphopantetheine arm on the ACP. ACPs are post translationally modified by 4'-phosphopantetheinyl transferases. These enzymes attach the phosphopantetheine arm from CoA onto a conserved serine residue in the ACP. This allows for a free thiol group that receives the acyl group from the AT. Once the acyl group is loaded onto the ACP, the KS catalyzes a Claisen-like condensation reaction between the starter unit on the KS and the acyl group on the ACP. The reaction results in a B-keto thioester. For example, if a PKS uses a malonyl group, each elongation step results in an addition of 2 carbons and 1 oxygen atom (Figure 2.9). After this step, the growing polyketide chain may be elongated further in the next module or undergo modification from additional tailoring domains in the module. The KR reduces the β-carbonyl into an alcohol. The DH can further reduce the alcohol by dehydration into α, β-unsaturated thioester. Finally, the ER can completely reduce the double bond to a single bond.[13] Along with the tailoring domains mentioned, methyl transferases can also incorporate methyl groups onto the growing polyketide chain.

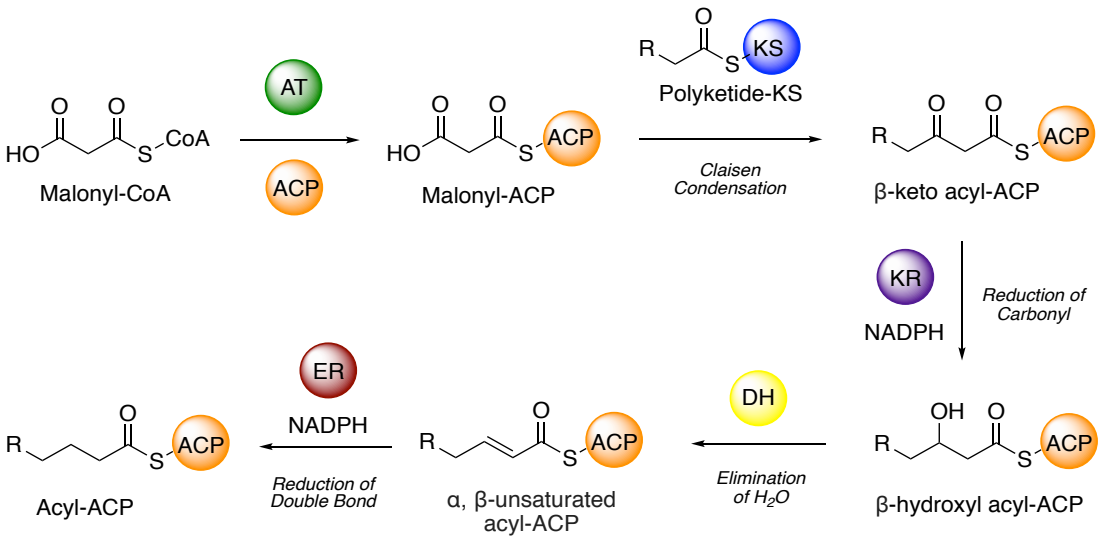


Figure 2.9. Domains in each module of type I polyketide synthases with malonyl-CoA.
1. The AT facilitates the transfer of the malonyl group from malonyl-CoA onto the ACP **2**. The KS catalyzes a Claisen-like condensation reaction between the malonyl group on the ACP and the growing polyketide chain on the KS **3**. The KR reduces the carbonyl at the β position to an alcohol **4**. The DH catalyzes the elimination of the hydroxyl producing a double bond **5**. The ER reduces the double bond into a single bond.

Once the elongation step is complete for that module, the process repeats for each module until it reaches a thioesterase (TE).[20] The TE cuts the polyketide chain off the phosphopantetheine arm. Note that in many trans-AT PKSs, there may be non-elongating modules. These non-elongating modules do not have a conserved histidine in the HTGTG motif in the KS and therefore cannot catalyze the condensation reaction. Instead these modules may result in no modification or a variety of modifications that do not involve elongation of the polyketide chain.[13] Although this is the basic foundation of

PKSs, some BGCs are more complex and may include additional enzymes downstream for further modification.

2.1.8 Determining the biosynthetic gene cluster of NP-34

In the genomic era, sequencing and genome mining have advanced the discovery of BGCs.[21,22] The first step in determining the BGC of NP-34 was to sequence the genome of SNC-034, also known as *Streptomyces malachitospinus*. The genome was sequenced and assembled by Genewiz. The assembled genome was then submitted to antiSMASH for secondary metabolite analysis (Figure 2.13). Since the tartrolon BGC from *Teredinibacter turnerae* is the only annotated BGC of any boronated polyketide macrolide, it was used as a reference for building our proposed BGC (Figure 2.14).[9]

2.1.8a Initial approaches to validating the BGC of NP-34

Proving the BGC was one of the most challenging aspects of this project. Heterologous expression is a common method to not only prove the BGC, but also increase production of secondary metabolites.[23,24] Initially, transformation associated recombination (TAR) was attempted in *Saccharomyces cerevisiae*.[25] What proved difficult was the large size of the BGC and how GC-rich *S. malachitospinus* is. Capture and assembly of the BGC was unsuccessful. The second approach was to use CRISPR/Cas9

directly in the organism SNC-034.[26] The goal was to knockout a gene necessary for NP-34 production. Unfortunately conjugating any plasmid into the organism proved unsuccessful. This could indicate this species is very difficult to genetically modify or genetically intractable. Lastly, the company Varigen Biosciences specializes in assembling BGCs in bacterial artificial chromosomes (BACs). They were successful at assembling a BAC containing the BGC by using CRISPR/Cas9 on gDNA extracted from SNC-034.

2.1.8b Heterologous expression of NP-34 in *Streptomyces albus J1074*

Because *Streptomyces coelicolor* M1154 has successfully been used as a heterologous host, it was initially chosen for conjugation.[27] Once the BAC was received, it was purified and conjugated into *S. coelicolor* M1154. Although the BGC was effectively conjugated into *S. coelicolor* M1154, as verified by PCR and sequencing, NP-34 was not detected by ^{11}B-NMR or HRMS. Based on RT-PCR it appeared the Trans-AT was not transcribing. Trans-AT was constructed on a plasmid with a constitutive promoter pKY01-ermE* and conjugated into *S. coelicolor* M1154 containing the BGC.[28] Extracts still showed no trace of NP-34. The BAC was conjugated into a new host, *S. albus J1074* and NP-34 was finally detected in organic extracts by ^{11}B-NMR and HRMS.[29]

2.2 Results

2.2.1 Confirmation of NP-34 in SNC-034 crude extracts

To confirm the presence of NP-34 in crude extracts of SNC-034, ^{11}B-NMR spectra were compared to purified samples of NP-34 (Figure 2.10). The crude extract had a sharp peak at approximately 10.5 p.pm., confirming the presence of NP-34 or a related boronated compound.

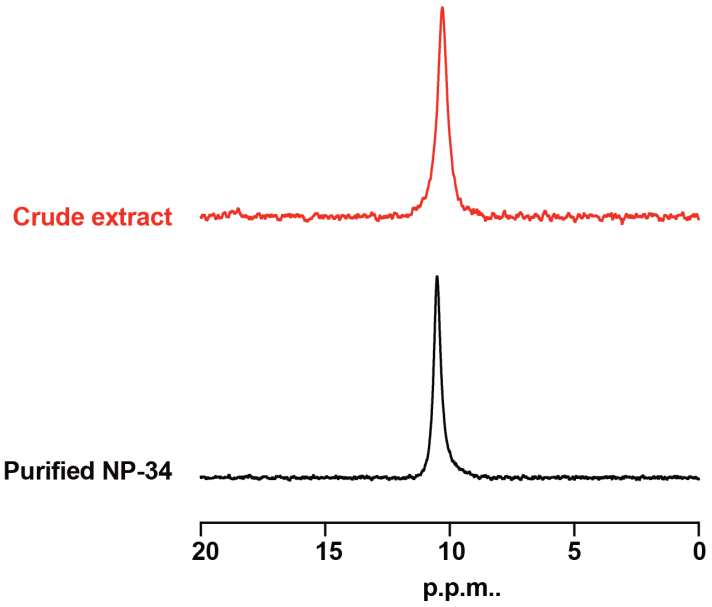


Figure 2.10. ^{11}B-NMR spectrum of purified NP-34 and crude organic extract from SNC-034.
The observed shift at approximately 10.5 ppm can be used to identify NP-34 in crude organic extracts of SNC-034.

To further verify the presence of NP-34 specifically, the crude extract was subjected to HRMS. NP-34 was present, at m/z = 845.45, with a distinct boron splitting pattern (Figure 2.11). In addition to NP-34, there were multiple analogs of NP-34 present in the crude extract of SNC-034 at m/z = 775.41,

817.42, 831.43, 859.42, 873.44, 887.44, 901.47, and 915.49. Figure 2 displays the molecular structure associated with the observed analogs of NP-34. One thing to note here is that the relative abundance of NP-34 is lower than some of these other analogs. This could be due to several factors such as growing conditions or the duration of growth.

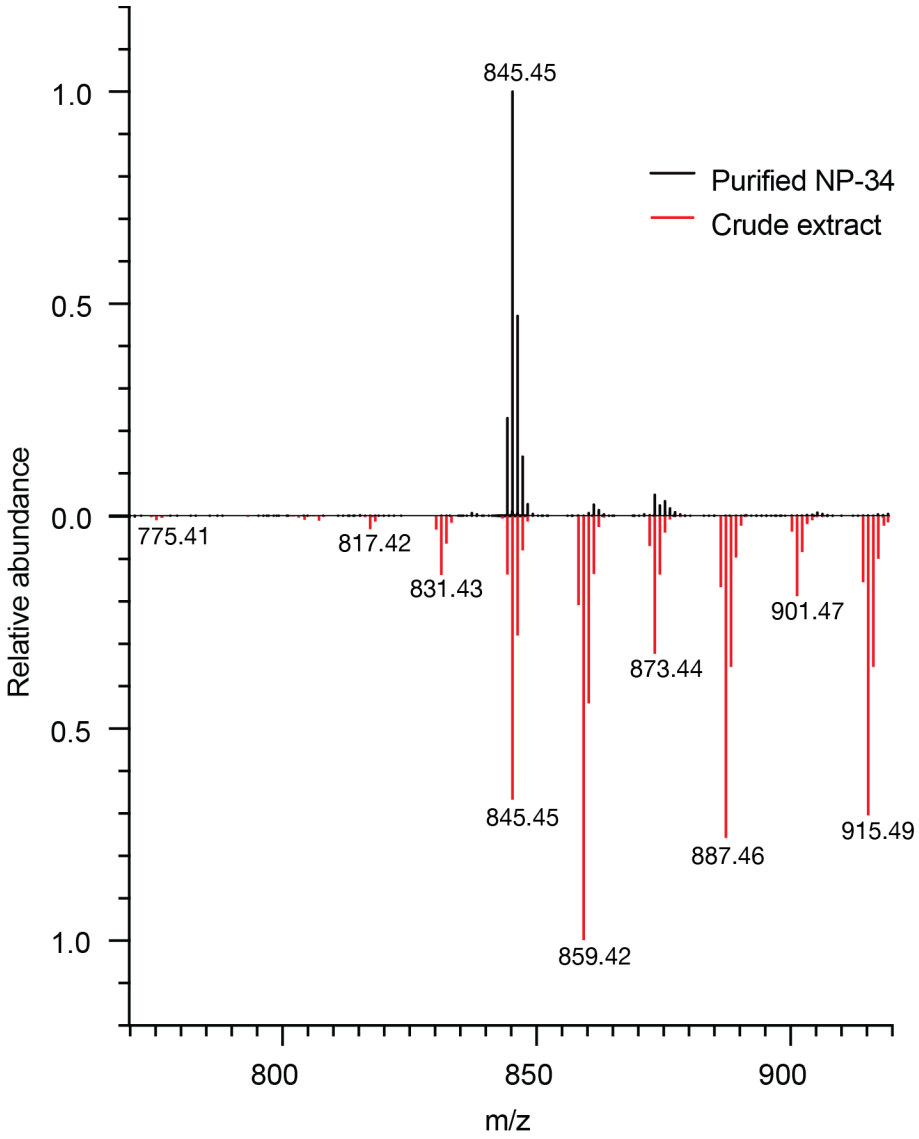


Figure 2.11. HRMS of purified NP-34 and crude organic extracts from SNC-034 in methanol.
Black: purified NP-34 from SNC-034 at m/z = 845.45 and Red: crude organic extract from SNC-034 with NP-34 analogs at m/z = 775.41, 817.42, 831.43, 845.45, 859.42, 873.44, 887.44, 901.47, and 915.49.

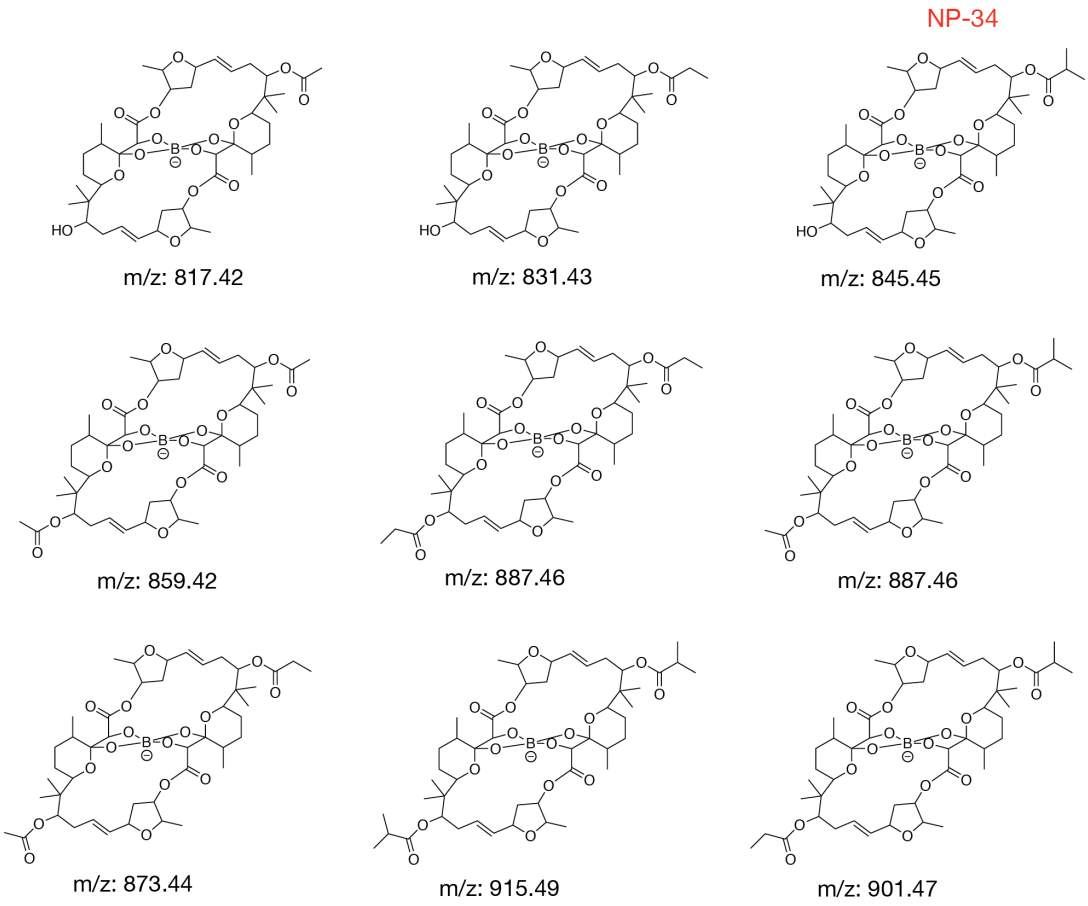
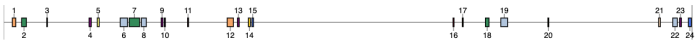


Figure 2.12 Molecular structures and associated m/z value of NP-34 analogs observed in SNC-034 extracts by HRMS in negative mode.

2.2.2 Sequencing and secondary metabolite analysis

antiSMASH analysis predicted 25 regions in the genome containing BGCs (Figure 2.13).[30] These regions included BGCs of PKSs, NRPS, RiPP-like, terpenes, lanthipeptides, thioamitides, butyrolactones, siderophores, LAPs, melanin, RRE-containing, thiopeptides, ectoines, NAPAAs, and indoles. Of those 23 regions, only 4 were predicted to be PKSs, and only one type I Trans-AT PKS. Since the BGC of tartrolons is known, it was easy to

predict the type I Trans-AT PKS belonged to NP-34.[9] Figure 2.14 compares the similarities and between each BGC.[9,30,31] Both BGCs contain all necessary genes for a trans-AT-PKS.



Region	Type	From	To	Most similar known cluster		Similarity
Region 1	T3PKS	100,034	140,544	germicidin	Other	100%
Region 2	NRP-metallophore, NRPS	208,333	263,692	coelichelin	NRP	81%
Region 3	RiPP-like	499,817	507,635	informatipeptin	RiPP:Lanthipeptide	42%
Region 4	terpene	990,857	1,016,571	hopene	Terpene	100%
Region 5	lanthipeptide-class-iii	1,086,138	1,108,747	SapB	RiPP:Lanthipeptide	100%
Region 6	transAT-PKS, thioamitides, butyrolactone	1,352,140	1,437,672	omnipeptin	NRP:Cyclic depsipeptide	9%
Region 7	NRPS	1,456,745	1,577,625	atratumycin	NRP	7%
Region 8	NRP-metallophore, NRPS, NI-siderophore	1,598,098	1,655,777	mirubactin	NRP	50%
Region 9	terpene	1,827,727	1,848,739	geosmin	Terpene	100%
Region 10	RiPP-like	1,865,632	1,876,219			
Region 11	NI-siderophore	2,134,372	2,145,420			
Region 12	T2PKS	2,593,775	2,666,347	spore pigment	Polyketide	66%
Region 13	terpene	2,713,227	2,734,036	albaflavenone	Terpene	100%
Region 14	LAP	2,839,805	2,862,972	oryzanaphthopyran A/oryzanaphthopyran B/oryzanaphthopyran C/oryzanthrone A/oryzanthrone B/chlororyzanthrone A/chlororyzanthrone B	Polyketide	6%
Region 15	RRE-containing	2,881,044	2,899,247	niphimycins C-E	Polyketide	6%
Region 16	NI-siderophore	5,215,415	5,227,184	desferrioxamin B/desferrioxamine E	Other	83%
Region 17	melanin	5,322,921	5,333,442	istamycin	Saccharide	4%
Region 18	NRPS-like	5,591,098	5,633,075	streptozotocin	Other	23%
Region 19	T3PKS, thiopeptide	5,770,142	5,849,677	granaticin	Polyketide:Type II polyketide	16%
Region 20	ectoine	6,316,507	6,326,905	ectoine	Other	100%
Region 21	indole	7,601,186	7,622,310	5-dimethylallylindole-3-acetonitrile	Other	100%
Region 22	terpene, NAPAA	7,765,831	7,818,346	isorenieratene	Terpene	100%
Region 23	terpene	7,845,255	7,866,316	2-methylisoborneol	Terpene	100%
Region 24	LAP, thiopeptide, RRE-containing	7,947,201	7,981,572	teicoplanin	NRP:Glycopeptide	9%

Figure 2.13 antiSMASH secondary metabolite BGC prediction. Predicted BGCs include PKSs, NRPS, RiPP-like, terpenes, lanthipeptides, thioamitides, butyrolactones, siderophores, LAPs, melanin, RRE-containing, thiopeptides, ectoines, NAPAAs, and indoles. [30] Region 6, the transAT-PKS, is the predicted BGC of NP-34.

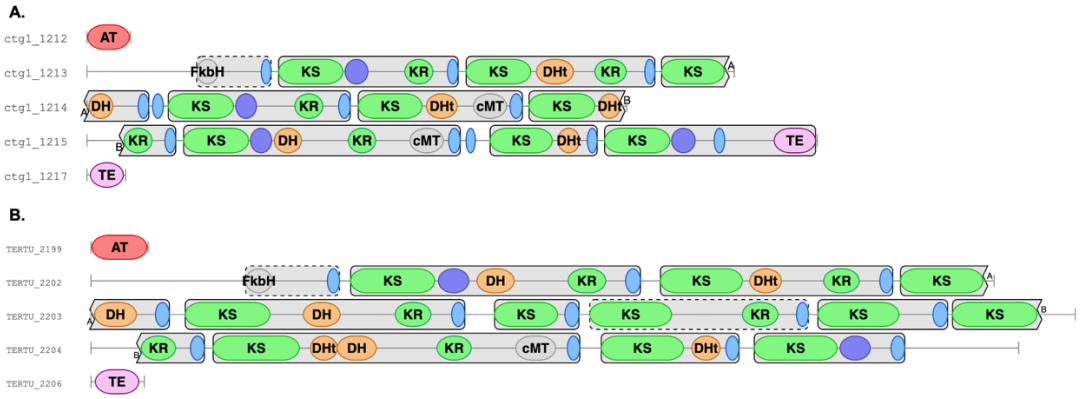


Figure 2.14. Comparison of antiSMASH prediction of the BGC at Region 6 of a transAT-PKS in SNC-034 and MiBIG analysis of the published tartrolon BGC.
AntiSMASH predicted BGC of a transAT-PKS.[30] (B) MiBIG results of published tartrolon BGC.[9,31] Both BGCs have similar features indicating the predicted BGC from antiSMASH may be responsible for producing NP-34. AT: acyltransferase, FkbH: unknown overall function, blue oval: acyl carrier protein (ACP), KS: ketosynthase, purple oval: AT docking site, KR: ketoreductase, DHt/DH: dehydratase, cMT: methyltransferase, TE: thioesterase

2.2.3 Proposed biosynthetic gene cluster of NP-34

After sequencing, genome mining, and literature review, the precited BGC of NP-34 is represented in Figure 2.15.[9,30–32] It is important to note this is a rough prediction, there is much more experimental analysis that needs to be done in order to fully establish the correct BGC. The starter unit is proposed to be lactate (green). The acyl group added during each Claisen-like condensation reaction was determined to be from malonyl-CoA. As stated in the introduction, each elongating module contains a KS and ACP. Since this is a trans-AT PKS, the AT moves between modules, transferring a malonyl group from malonyl-CoA onto the ACP. Tailoring domains include

KR, DH/DHt, and MT. After addition of the acyl group, each KR reduces the β-carbonyl into an alcohol and the DH/DHt reduces the hydroxyl into an α, β-unsaturated thioester if either are present. The methyl transferase adds 1-2 methyl groups on the α-carbon. The growing polyketide chains passes through each module until it reaches the TE domain. There are two TEs present in the BGC. It is still unclear which TE is doing which function. It is hypothesized the TEs are involved in cutting the polyketide off the phosphopantetheine arm and facilitating with dimerization and cyclization of the polyketide macrolide. The monooxygenase is predicted to add a hydroxyl group onto the α-carbon. Boron is suspected to non-enzymatically form a $MB(OR)_4$ complex from 4 hydroxyls. Finally, the acyl group is added by the acyltranferase to give NP-34.

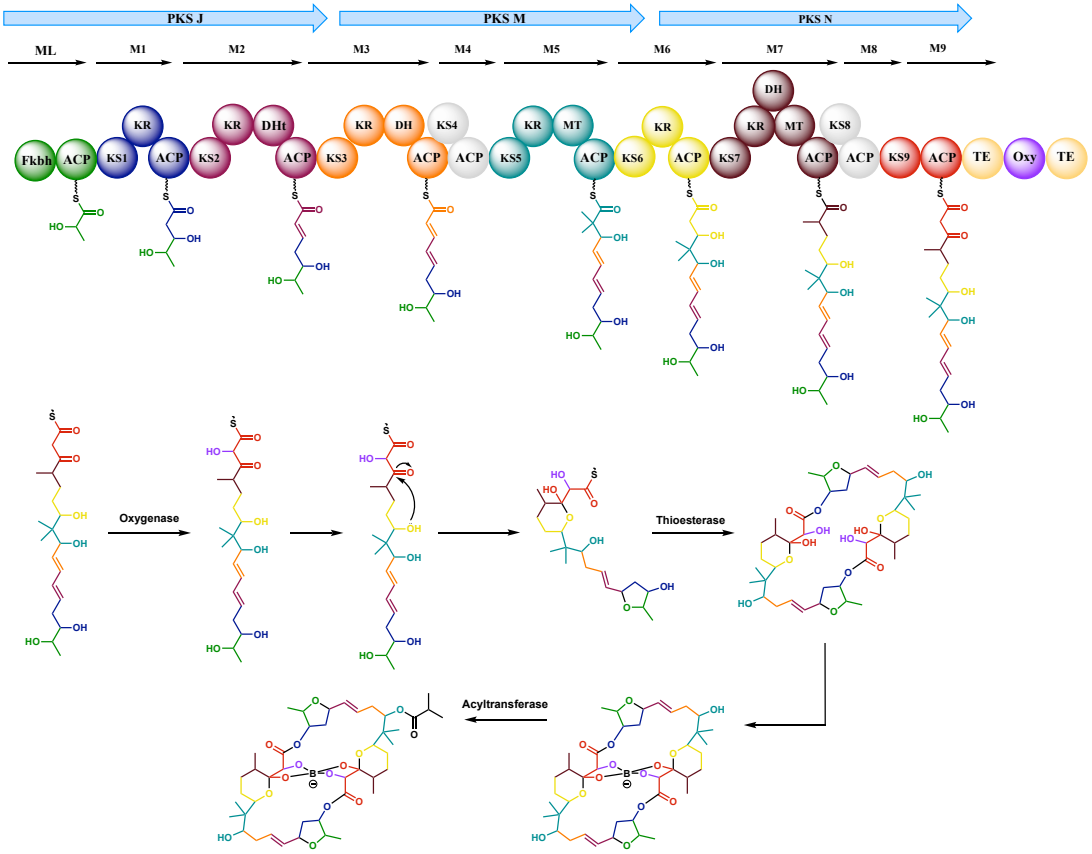


Figure 2.15. Proposed biosynthetic gene cluster of NP-34 produced by *Streptomyces malachitospinus*.
Each blue arrow indicates the 3 PKS genes J, M, and N. ML: module loading, M1-M9: modules 1-9, AT: acyltransferase, FkbH: unknown overall function, ACP: acyl carrier protein, KS: ketosynthase, KR: ketoreductase, DHt/DH: dehydratase, MT: methyltransferase, TE: thioesterase, and Oxy: oxygenase.

It is still unclear which domains are non-elongating KS° domains (grey) and if they are involved in ring formation and/or influence stereochemistry. Stereochemistry is not shown for this reason. antiSMASH predicted KS7 in module 7 to be non-elongating, however, further analysis of AlphaFold2 structures (Figure 2.16) revealed an intact active site for KS7, contradicting

these antiSMASH results.[30,33] Further analysis should be done to accurately confirm this prediction as well as the function of each TE mentioned above.

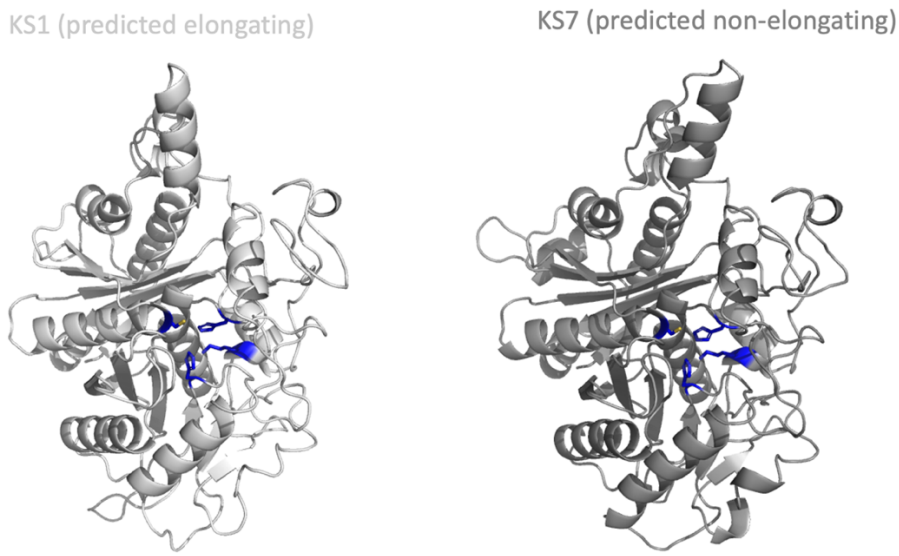


Figure 2.16. Predicted AlphaFold structures of KS domains.
Predicted structures of the KS1 (left) and KS7 (right) catalytic domains. The active site tetrad of residues (blue) are conserved in both KS domains.

2.2.4 Heterologous expression of NP-34 in *Streptomyces albus J1074*

PCR, Sanger sequencing, ^{11}B-NMR, and HRMS was used to confirm that the BGC of NP-34 was successfully conjugated into *S. albus* J1074. The core genes of the BGC coding for Sfp, PksE, Trans-AT, KS1, KS5, TE, OX, LuxR, and AT were all first confirmed by PCR and Sanger sequencing. Colonies with the correct genes were upscaled to 1 L growths and the organic material was extracted from the cultures. Extracts were screened by ^{11}B-NMR confirming the presence of a peak at approximately 10.5 p.p.m., suggesting

that heterologous expression of the NP-34 BAC did produce a boronated compound (Figure 2.17).

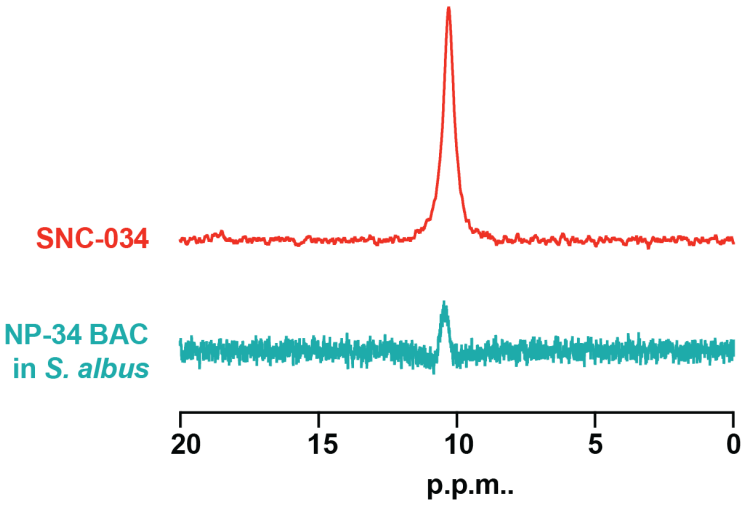


Figure 2.17 ^{11}B-NMR of crude organic extracts from SNC-034 and *S. albus* J1074 expressing NP-34.
SNC-034 (red) with a shift at approximately 10.5 ppm with the same shift observed in *S. albus* J1074 expressing the NP-34 BAC.

Further analysis by HRMS revealed the presence of each of the NP-34 analogs observed in SNC-034 extracts (Figure 2.18). Interestingly, the relative abundance of NP-34 is the highest of the observed analogs present in extracts from *S. albus* J1074 heterologously expressing the NP-34 BAC, as opposed to SNC-034. This could be due to growth conditions, or possibly cellular conditions and/or regulatory elements that differ between SNC-034 and *S. albus* J1074.

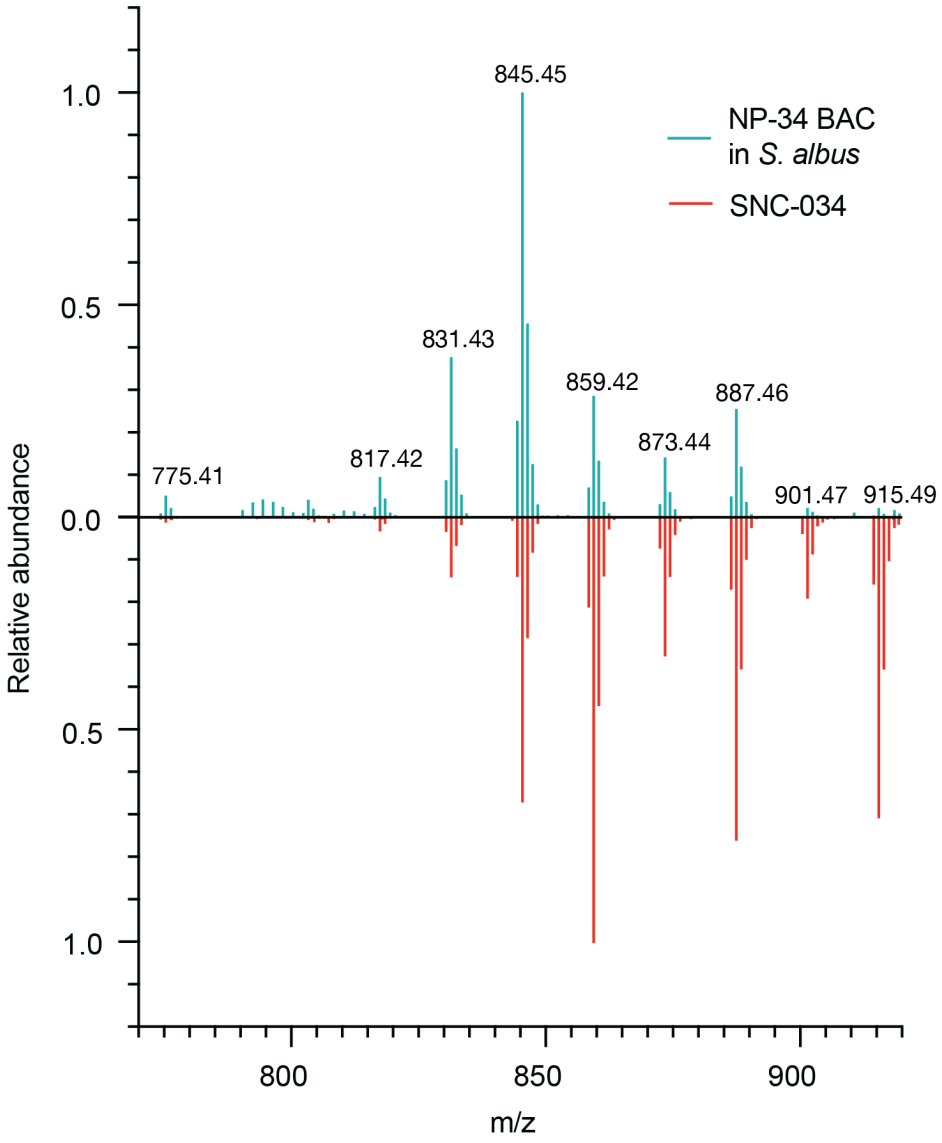


Figure 2.18 HRMS comparison of NP-34 and analogs in SNC-034 and the NP-34 BAC expressed in *S. albus J1074*.
The same masses were observed in both SNC-034 (red) and *S. albus* J1074 expressing the NP-34 BAC. The NP-34 analogs include m/z = 775.41, 817.42, 831.43, 845.45, 859.42, 873.44, 887.44, 901.47, and 915.49.

2.3 Discussion

After sequencing of SNC-034 (*S. malachitospinus*), genome mining, and heterologous expression of the BGC in *S. albus* J1074, the set of genes involved in the BGC that produces NP-34 were determined conclusively. That being said, it should be noted that the exact mechanism by which NP-34 is produced by these enzymes is still not fully known. Much more work needs to be done to elucidate the molecular mechanisms involved in each step of NP-34 synthesis, such as the mechanism by which the rings form and how the non-elongating KS$^{\circ}$ domains are formed. Recent advances in bioinformatics and protein structure prediction will help further our understanding of polyketide synthases and how this important cellular machinery leads to the production of natural products that could benefit human health.

Selective insecticides are extremely important resources that could aid in the prevention of disease. The discovery and characterization of NP-34 as a selective mosquitocide provides a natural product-based framework to develop similar molecules that improve selectivity and potency against mosquitos. Importantly, the results of this study also show that the BGC that produces NP-34, and other potentially useful analogs, can be heterologously expressed in an alternate organism allowing for genetic control of biosynthetic products.

2.4 References

(1) Dias, D. A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012**, *2* (2), 303–336. https://doi.org/10.3390/metabo2020303.

(2) *NOAA Ocean Explorer: Deep Sea Medicines 2003: Products*. https://oceanexplorer.noaa.gov/explorations/03bio/background/products/products.html (accessed 2023-11-06).

(3) Maplestone, R. A.; Stone, M. J.; Williams, D. H. The Evolutionary Role of Secondary Metabolites — a Review. *Gene* **1992**, *115* (1), 151–157. https://doi.org/10.1016/0378-1119(92)90553-2.

(4) Watve, M. G.; Tickoo, R.; Jog, M. M.; Bhole, B. D. How Many Antibiotics Are Produced by the Genus Streptomyces? *Arch. Microbiol.* **2001**, *176* (5), 386–390. https://doi.org/10.1007/s002030100345.

(5) Clardy, J.; Fischbach, M. A.; Currie, C. R. The Natural History of Antibiotics. *Curr. Biol.* **2009**, *19* (11), R437–R441. https://doi.org/10.1016/j.cub.2009.04.001.

(6) Duke, S. O.; Cantrell, C. L.; Meepagala, K. M.; Wedge, D. E.; Tabanca, N.; Schrader, K. K. Natural Toxins for Use in Pest Management. *Toxins* **2010**, *2* (8), 1943–1962. https://doi.org/10.3390/toxins2081943.

(7) O'Neal, M. A.; Posner, B. A.; Coates, C. J.; Abrams, J. M. A Cell-Based Screening Platform Identifies Novel Mosquitocidal Toxins. *J. Biomol. Screen.* **2013**, *18* (6), 688–694. https://doi.org/10.1177/1087057113476952.

(8) DeFrancesco, H.; Dudley, J.; Coca, A. Boron Chemistry: An Overview. In *Boron Reagents in Synthesis*; ACS Symposium Series; American Chemical Society, 2016; Vol. 1236, pp 1–25. https://doi.org/10.1021/bk-2016-1236.ch001.

(9) Elshahawi, S. I.; Trindade-Silva, A. E.; Hanora, A.; Han, A. W.; Flores, M. S.; Vizzoni, V.; Schrago, C. G.; Soares, C. A.; Concepcion, G. P.; Distel, D. L.; Schmidt, E. W.; Haygood, M. G. Boronated Tartrolon Antibiotic Produced by Symbiotic Cellulose-Degrading Bacteria in Shipworm Gills. *Proc. Natl. Acad. Sci.* **2013**, *110* (4). https://doi.org/10.1073/pnas.1213892110.

(10) Shimizu, Y.; Ogasawara, Y.; Matsumoto, A.; Dairi, T. Aplasmomycin and Boromycin Are Specific Inhibitors of the Futalosine Pathway. *J. Antibiot. (Tokyo)* **2018**, *71* (11), 968–970. https://doi.org/10.1038/s41429-018-0087-2.

(11) Hemscheidt, T.; Puglisi, M. P.; Larsen, L. K.; Patterson, G. M. L.; Moore, R. E.; Rios, J. L.; Clardy, J. Structure and Biosynthesis of Borophycin, a New Boeseken Complex of Boric Acid from a Marine Strain of the Blue-Green Alga Nostoc Linckia. *J. Org. Chem.* **1994**, *59* (12), 3467–3471. https://doi.org/10.1021/jo00091a042.

(12) Laurent, D.; Pietra, F. Antiplasmodial Marine Natural Products in the Perspective of Current Chemotherapy and Prevention of Malaria. A Review.

Mar. Biotechnol. **2006**, *8* (5), 433–447. https://doi.org/10.1007/s10126-006-6100-y.

(13) N. Helfrich, E. J.; Piel, J. Biosynthesis of Polyketides by Trans -AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33* (2), 231–316. https://doi.org/10.1039/C5NP00125K.

(14) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Curr. Opin. Chem. Biol.* **2003**, *7* (2), 285–295. https://doi.org/10.1016/S1367-5931(03)00020-6.

(15) Khosla, C. Harnessing the Biosynthetic Potential of Modular Polyketide Synthases. *Chem. Rev.* **1997**, *97* (7), 2577–2590. https://doi.org/10.1021/cr960027u.

(16) Khosla, C. Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem.* **2009**, *74* (17), 6416–6420. https://doi.org/10.1021/jo9012089.

(17) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Nat Prod Rep* **2007**, *24* (1), 162–190. https://doi.org/10.1039/B507395M.

(18) B. Austin, M.; P. Noel, J. The Chalcone Synthase Superfamily of Type III Polyketide Synthases. *Nat. Prod. Rep.* **2003**, *20* (1), 79–110. https://doi.org/10.1039/B100917F.

(19) Katsuyama, Y.; Ohnishi, Y. Chapter Sixteen - Type III Polyketide Synthases in Microorganisms. In *Methods in Enzymology*; Hopwood, D. A., Ed.; Natural Product Biosynthesis by Microorganisms and Plants, Part A;

Academic Press, 2012; Vol. 515, pp 359–377. https://doi.org/10.1016/B978-0-12-394290-6.00017-3.

(20) Zhou, Y.; Tao, W.; Qi, Z.; Wei, J.; Shi, T.; Kang, Q.; Zheng, J.; Zhao, Y.; Bai, L. Structural and Mechanistic Insights into Chain Release of the Polyene PKS Thioesterase Domain. *ACS Catal.* **2022**, *12* (1), 762–776. https://doi.org/10.1021/acscatal.1c04991.

(21) Albarano, L.; Esposito, R.; Ruocco, N.; Costantini, M. Genome Mining as New Challenge in Natural Products Discovery. *Mar. Drugs* **2020**, *18* (4), 199. https://doi.org/10.3390/md18040199.

(22) Luo, Y.; Cobb, R. E.; Zhao, H. Recent Advances in Natural Product Discovery. *Curr. Opin. Biotechnol.* **2014**, *0*, 230–237. https://doi.org/10.1016/j.copbio.2014.09.002.

(23) Jia Zhang, J.; Tang, X.; S. Moore, B. Genetic Platforms for Heterologous Expression of Microbial Natural Products. *Nat. Prod. Rep.* **2019**, *36* (9), 1313–1332. https://doi.org/10.1039/C9NP00025A.

(24) Huo, L.; J. Hug, J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R. Heterologous Expression of Bacterial Natural Product Biosynthetic Pathways. *Nat. Prod. Rep.* **2019**, *36* (10), 1412–1436. https://doi.org/10.1039/C8NP00091C.

(25) Tang, X.; Li, J.; Millán-Aguiñaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.; Jensen, P. R.; Mantovani, S. M.; Moore, B. S. Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-Directed

Genome Mining. *ACS Chem. Biol.* **2015**, *10* (12), 2841–2849. https://doi.org/10.1021/acschembio.5b00658.

(26) Cobb, R. E.; Wang, Y.; Zhao, H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. *ACS Synth. Biol.* **2015**, *4* (6), 723–728. https://doi.org/10.1021/sb500351f.

(27) Gomez-Escribano, J. P.; Bibb, M. J. Engineering Streptomyces Coelicolor for Heterologous Expression of Secondary Metabolite Gene Clusters. *Microb. Biotechnol.* **2011**, *4* (2), 207–215. https://doi.org/10.1111/j.1751-7915.2010.00219.x.

(28) Bauman, K. D.; Li, J.; Murata, K.; Mantovani, S. M.; Dahesh, S.; Nizet, V.; Luhavaya, H.; Moore, B. S. Refactoring the Cryptic Streptophenazine Biosynthetic Gene Cluster Unites Phenazine, Polyketide, and Nonribosomal Peptide Biochemistry. *Cell Chem. Biol.* **2019**, *26* (5), 724-736.e7. https://doi.org/10.1016/j.chembiol.2019.02.004.

(29) Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. Insights into Naturally Minimised Streptomyces Albus J1074 Genome. *BMC Genomics* **2014**, *15* (1), 97. https://doi.org/10.1186/1471-2164-15-97.

(30) Kai Blin, Simon Shaw, Hannah E Augustijn, Zachary L Reitz, Friederike Biermann, Mohammad Alanjary, Artem Fetter, Barbara R Terlouw, William W. Metcalf, Eric JN Helfrich, Gilles P van Wezel, Marnix. H Medema & Tilmann Weber. antiSMASH 7.0: New and Improved Predictions for

Detection, Regulation, Chemical Structures, and Visualisation. *Nucleic Acids Res.* **2023**. https://doi.org/doi: 10.1093/nar/gkad344.

(31) Terlouw, B. R.; Blin, K.; Navarro-Muñoz, J. C.; Avalon, N. E.; Chevrette, M. G.; Egbert, S.; Lee, S.; Meijer, D.; Recchia, M. J. J.; Reitz, Z. L.; van Santen, J. A.; Selem-Mojica, N.; Tørring, T.; Zaroubi, L.; Alanjary, M.; Aleti, G.; Aguilar, C.; Al-Salihi, S. A. A.; Augustijn, H. E.; Avelar-Rivas, J. A.; Avitia-Domínguez, L. A.; Barona-Gómez, F.; Bernaldo-Agüero, J.; Bielinski, V. A.; Biermann, F.; Booth, T. J.; Carrion Bravo, V. J.; Castelo-Branco, R.; Chagas, F. O.; Cruz-Morales, P.; Du, C.; Duncan, K. R.; Gavriilidou, A.; Gayrard, D.; Gutiérrez-García, K.; Haslinger, K.; Helfrich, E. J. N.; van der Hooft, J. J. J.; Jati, A. P.; Kalkreuter, E.; Kalyvas, N.; Kang, K. B.; Kautsar, S.; Kim, W.; Kunjapur, A. M.; Li, Y.-X.; Lin, G.-M.; Loureiro, C.; Louwen, J. J. R.; Louwen, N. L. L.; Lund, G.; Parra, J.; Philmus, B.; Pourmohsenin, B.; Pronk, L. J. U.; Rego, A.; Rex, D. A. B.; Robinson, S.; Rosas-Becerra, L. R.; Roxborough, E. T.; Schorn, M. A.; Scobie, D. J.; Singh, K. S.; Sokolova, N.; Tang, X.; Udwary, D.; Vigneshwari, A.; Vind, K.; Vromans, S. P. J. M.; Waschulin, V.; Williams, S. E.; Winter, J. M.; Witte, T. E.; Xie, H.; Yang, D.; Yu, J.; Zdouc, M.; Zhong, Z.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. MIBiG 3.0: A Community-Driven Effort to Annotate Experimentally Validated Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2023**, *51* (D1), D603–D610. https://doi.org/10.1093/nar/gkac1049.

(32) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

(33) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

Chapter 3

The effect of gene knockouts on the biosynthetic gene cluster of NP-34

3.1 Introduction

3.1.1 Gene knockouts

Gene knockouts, developed in the 1980s by Dr. Martin Evans, Dr. Oliver Smithies, and Dr. Mario Capecchi, have made a significant impact on understanding how genes function.[1,2] These gene knockouts were originally produced by homologous recombination. Homologous recombination occurs naturally in bacteria repairing damaged DNA and incorporating new genetic variability.[3,4,5] The RecA protein, involved in homologous recombination, is responsible for binding to single stranded DNA and searching for a homologous DNA sequence. Once a homologous complementary DNA sequence is found, a strand exchange reaction occurs. The donor DNA complimentary sequence binds to the single stranded DNA creating a new double stranded DNA complex.[5] This can either result in repair of the original sequence or the incorporation of a new DNA sequences. Scientists like Dr. Evans, Smithies, and Capecchi harnessed this naturally occurring function by deleting or changing sequences to better understand how genes function. Today there are multiple methods to create gene knockouts and knockdowns using techniques such as CRISPR/Cas9, RNAi, and TALENs.[6,7,8]

3.1.2 Design of NP-34 knockouts by homologous recombination

One of the major factors that was used to determine which knockout method to use depended on downstream screening methods for verifying the desired knockouts. Screening for knockouts can be tedious and time consuming due to techniques such as FACS (fluorescent assisted cell sorting) to isolate or enrich for cells with the desired knockout, followed by sequence validation. If the efficiency is low, this can take days if not weeks. CRISPR/Cas9, for example, can also lead to in-frame deletions that retain some of the genetic function of the gene to be knocked out, and therefor may require a careful targeting strategy. In the case of this study, homologous recombination was the method chosen to generate gene knockouts within the NP-34 BGC because it offered a straightforward selection technique by simultaneously knocking out the desired gene and inserting an antibiotic resistance gene.

To elucidate and confirm various enzymes in the NP-34 pathway a set of gene knockouts were created. These knockouts targeted a thioesterase, acyltransferase, and two LuxR regulatory genes (Figure 3.1). The reason these genes were chosen is because they could elucidate how the final steps or regulation of the final steps of the biosynthesis takes place. Knocking out sections of the core biosynthetic genes such as the Trans-AT and PKS will most likely disrupt protein-protein interactions leading to loss of the entire biosynthesis.

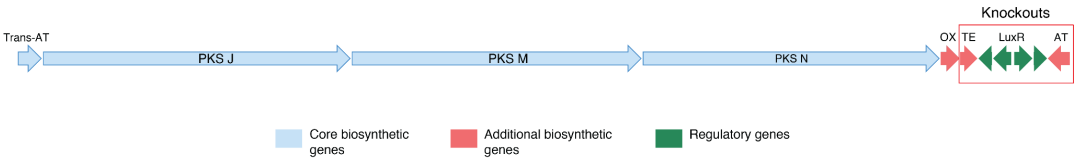


Figure 3.1. Targeted knockouts in NP-34 gene cluster
Core biosynthetic genes include Trans-AT (trans-acyltransferase) and PKS J, M, N (polyketide synthases). Additional biosynthetic genes include OX (monooxygenase), TE (thioesterase), and AT (acyltransferase). Regulatory genes include transcription factor LuxR. Targeted knockouts include TE, LuxR, and AT.

There are two acyltransferases (ATs) in the BGC. The first AT is thought to be involved in building the core structure of NP-34 by facilitating the PKS.[9] The AT catalyzes the transfer of a malonyl group from malonyl-CoA onto the ACP. The KS catalyzes a Claisen-like condensation reaction between the growing polyketide chain and malonyl group from the ACP.[10] The AT repeats this process for each module in the PKS. The downstream AT is hypothesized to control the last step of NP-34 production (Figure 3.4). The acyltransferase is thought to be responsible for transferring different acyl groups onto a hydroxyl group from different acyl-CoAs.

According to antiSMASH and BLAST, the downstream AT is predicted to be from the acyltransferase 3 superfamily.[11,12] Further structural analysis from Phyre2, predicted the AT to be a membrane bound protein.[13] The initial approach to test the hypothesis that the downstream AT is involved in the last step of NP-34 production, was to overexpress the AT protein separately and show that the precursor aplasmomycin can be acylated by the recombinant AT with different acyl groups. Unfortunately, expression of the AT was never

observed after multiple vectors, tags, cell types, growing conditions, and purifications were explored. Alternatively, since the NP-34 BAC was able to be expressed heterologously, knockout of the downstream AT could address these questions.

Thioesterases (TE) from type I PKS are generally found after the last module of a PKS and are involved in chain release and cyclization of the macrolide[14]. Knockout of the TE should prevent the growing polyketide chain from being cut off the ACP and dimerizing. Something interesting to note is there are two TEs present within the NP-34 BGC. One at the end of PKS N gene and one directly after the monooxygenase. The repetition could indicate the second TE is acting as a checking mechanism to ensure the chain is correctly cyclized or in the correct orientation before releasing it.[14] The TE knocked out in this study was the downstream TE directly after the monooxygenase.

LuxR is a regulatory protein or more specifically a transcription factor. Transcription factors are involved in activating or repressing genes.[15] LuxR can be both a transcriptional activator and repressor.[16] If LuxR is a transcriptional repressor, NP-34 biosynthesis should be upregulated if LuxR is knocked out. Since an overarching goal of this project was to increase production of NP-34, upregulation of NP-34 would be highly favored and therefore a valuable target to knockout. There are 4 regulatory genes that are suspected to be involved in regulation that are located between the

thioesterase and acyltransferase. According to antiSMASH, these genes are suspected to be a LuxR family DNA-binding response regulator, sensor histidine kinase, response regulator, and an additional sensor histidine kinase[11]. Since there was sequence overlap between the first two genes, both of those were knockout and replaced with AmpR.

3.1.3 Homologous recombination with pREDCas9

To generate these knockouts, homologous recombination was utilized using the pREDCas9 plasmid containing the lambda red system purchased from Addgene (Plasmid # 71541).[17] pREDCas9 contains useful genes for homologous recombination including RecA and the lambda red genes gam, exo, and beta. RecA is responsible for locating homologous sequences and exchanging DNA strands for recombination.[5] The lambda red gene gam protects the introduced drug cassette from being destroyed by exonucleases. Exo is an exonuclease that cuts back DNA 5'-3' to create a single stranded overhang or product. Beta protects the remaining single-stranded DNA and promotes annealing to the complementary sequence, in this case, the drug cassette containing the homologous sequence.[17,18] CRISPR/Cas9 can also be used to create double-stranded breaks in conjunction with the plasmid pGRB, which can be designed to contain a specific guide for targeted cutting.[6,17] The CRISPR/Cas9 component was not necessary to generate gene knockouts in this study, therefore it was not utilized.

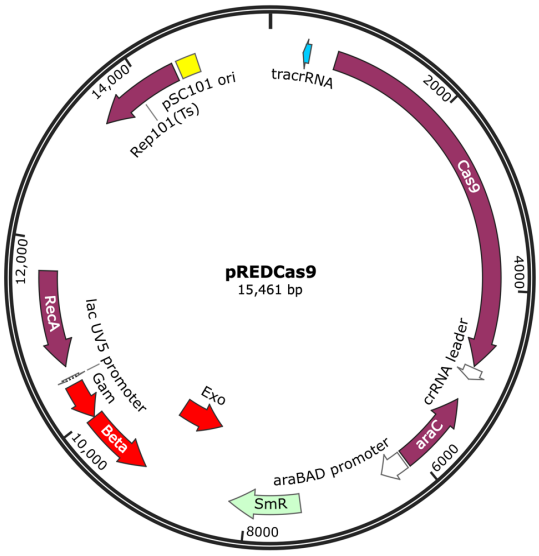


Figure 3.2. pREDCas9 plasmid utilized during homologous recombination
The pREDCas9 plasmid contains RecA, gam, beta, and exo genes to facilitate homologous recombination. SmR (aminoglycoside adenyltransferase) confers resistance to spectinomycin and streptomycin. Cas9 is an endonuclease that creates double-stranded breaks in conjunction with tracrRNA and crRNA. AraC is L-arabinose regulatory protein which is controlled by the araBAD promoter. pSC101 ori is a low copy origin of replication for Rep101(Ts), a temperature sensitive RepA protein that's required for replication of the plasmid.[17]

Drug cassettes containing the ampicillin resistance gene (AmpR) were used to replace the targeted genes TE, LuxR, and AT. To build the AmpR drug cassette, 70-bp primers were designed with 50-bp homologous to the upstream of the target sequence and 20- bp homologous to the beginning of the AmpR gene. This process was repeated for the reverse primer downstream of the target. The drug cassettes were generated by PCR (Figure 3.3). The NP-34 BAC was initially transformed into NEB Turbo *E. coli*

cells followed by the pREDCas9 plasmid. The drug cassette was then electroporated after induction of the pREDCas9 plasmid. Cells were plated on LB ampicillin plates and colonies were tested through PCR and sequencing. Once knockouts were confirmed, the modified BAC was conjugated into *S. albus* J1074. Samples were tested by ^{11}B-NMR and HRMS.

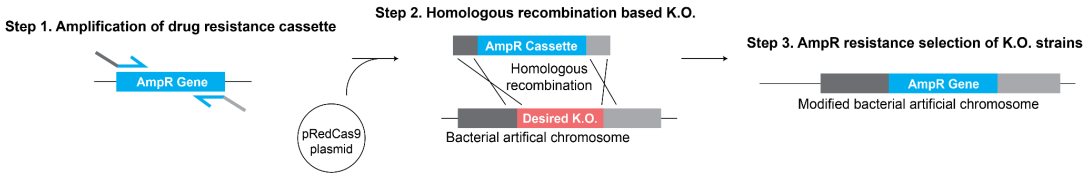


Figure 3.3. Knockouts produced by homologous recombination
Step 1: Drug cassettes were constructed by PCR amplified using an AmpR gene template. **Step 2**: pREDCas9 was transformed into *E. coli* containing the NP-34 BGC. The drug cassette was electroporated into E. coli containing the NP-34 BGC and pREDCas9 to initiate homologous recombination. **Step 3**: Knockouts were plated on ampicillin plates for selection.

3.2 Results

3.2.1 Analysis of analogs produced by *S. albus* J1074 expressing NP-34

Organic extracts from *S. albus* J1074 expressing the NP-34 BAC were analyzed by HRMS. While heterologous expression of the NP-34 BAC in this case does seems to preferentially produce NP-34, the relative abundance of alternate analogs suggests some level of promiscuity by the AT. A characteristic boron splitting pattern was observed for the different analogs of NP-34 with the most abundant analog NP-34 at m/z = 845.45 followed by m/z = 831.43, 859.42, 887.44, 873.44, 817.42, 775.41, 901.47, and 915.49 (Figure 3.4).

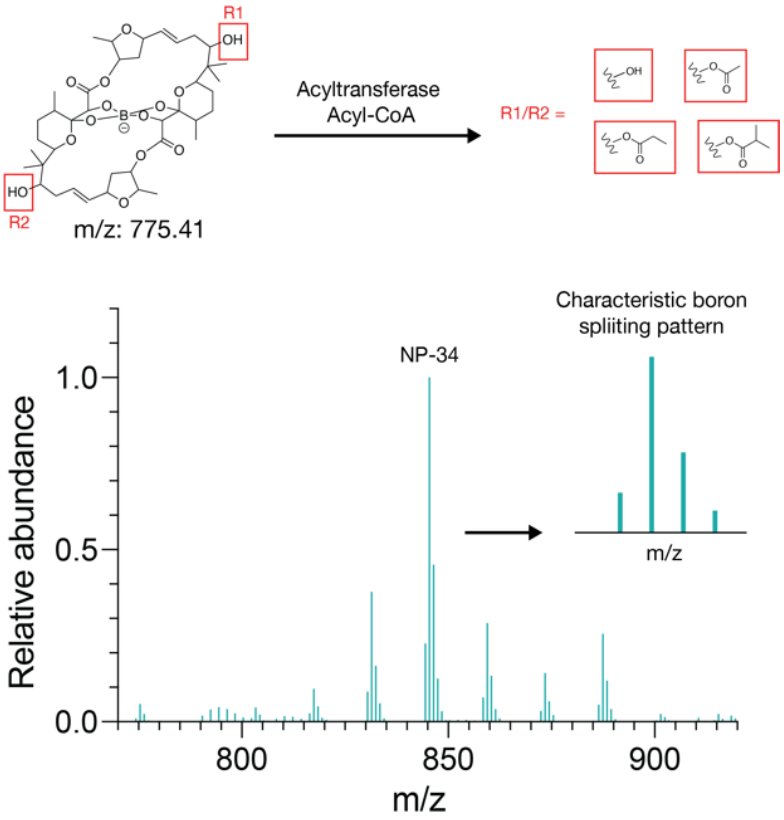


Figure 3.4. HRMS spectra of the proposed BGC of NP-34 BGC expressed in *S. albus* J1074 in negative mode.
The scheme above illustrates the analogs generated after the last step in the NP-34 biosynthesis with the addition of an acyl group on the hydroxyl in the heterologously expressed NP-34 BGC in *S. albus* J1074. **Top:** The precursor of NP-34 at m/z = 775.41 transfers 3 possible acyl groups. **Bottom:** The relative abundance of each analog and precursor after the last step in the NP-34 biosynthesis at m/z = 845.45, 831.43, 859.42, 887.44, 873.44, 817.42, 901.47, and 915.49.

Since the proposed precursor molecule at m/z = 775.41 is a dimer, there are two possible locations that the AT could add acyl groups resulting in a set of 9 possible analogs (Figure 3.5). Two of these analogs at m/z = 887.46 share the same mass but have different acyl groups attached on either side of the molecule.

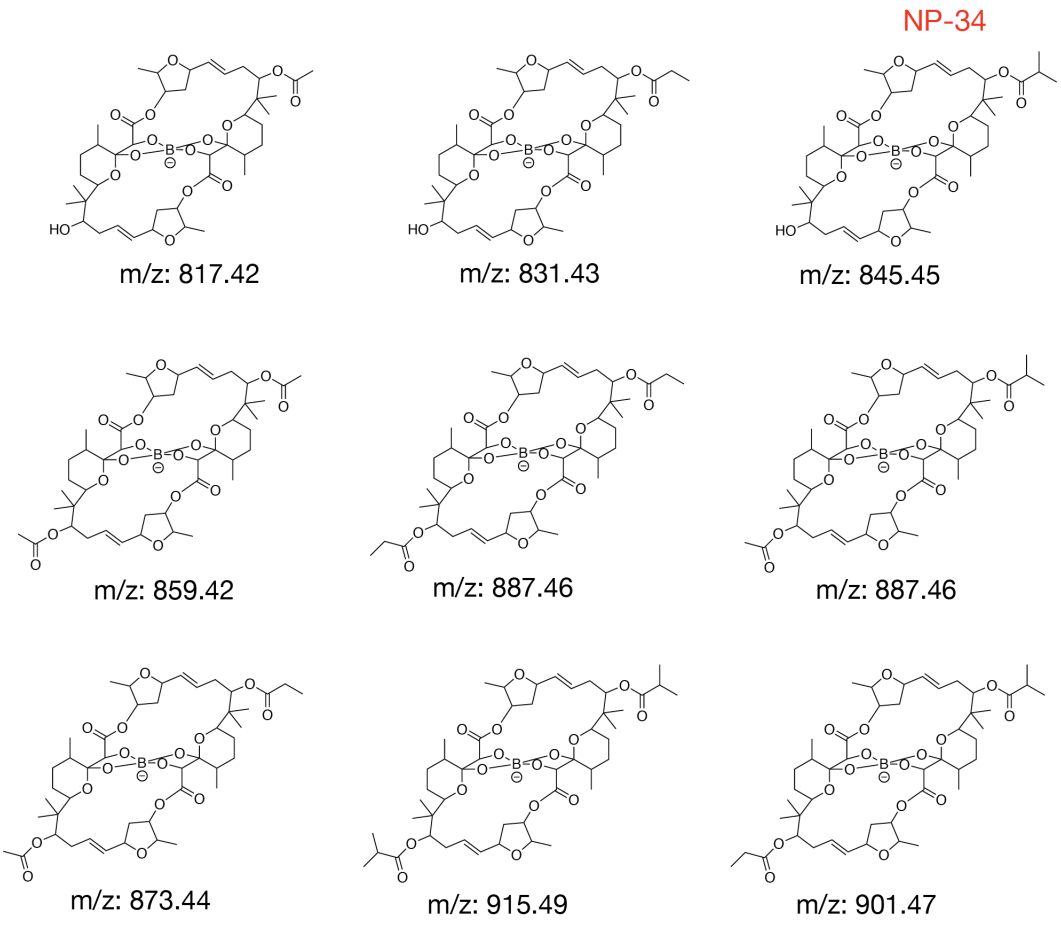


Figure 3.5. Molecular structures of the proposed analogs produced by the BGC of NP-34.
Each molecular structure corresponds to one of the acylated products produced by the downstream AT. Since the precursors molecule (m/z = 775.41) is proposed to be a dimer, acylation can occur on either side of the molecule resulting in a total of 9 possible analogs.

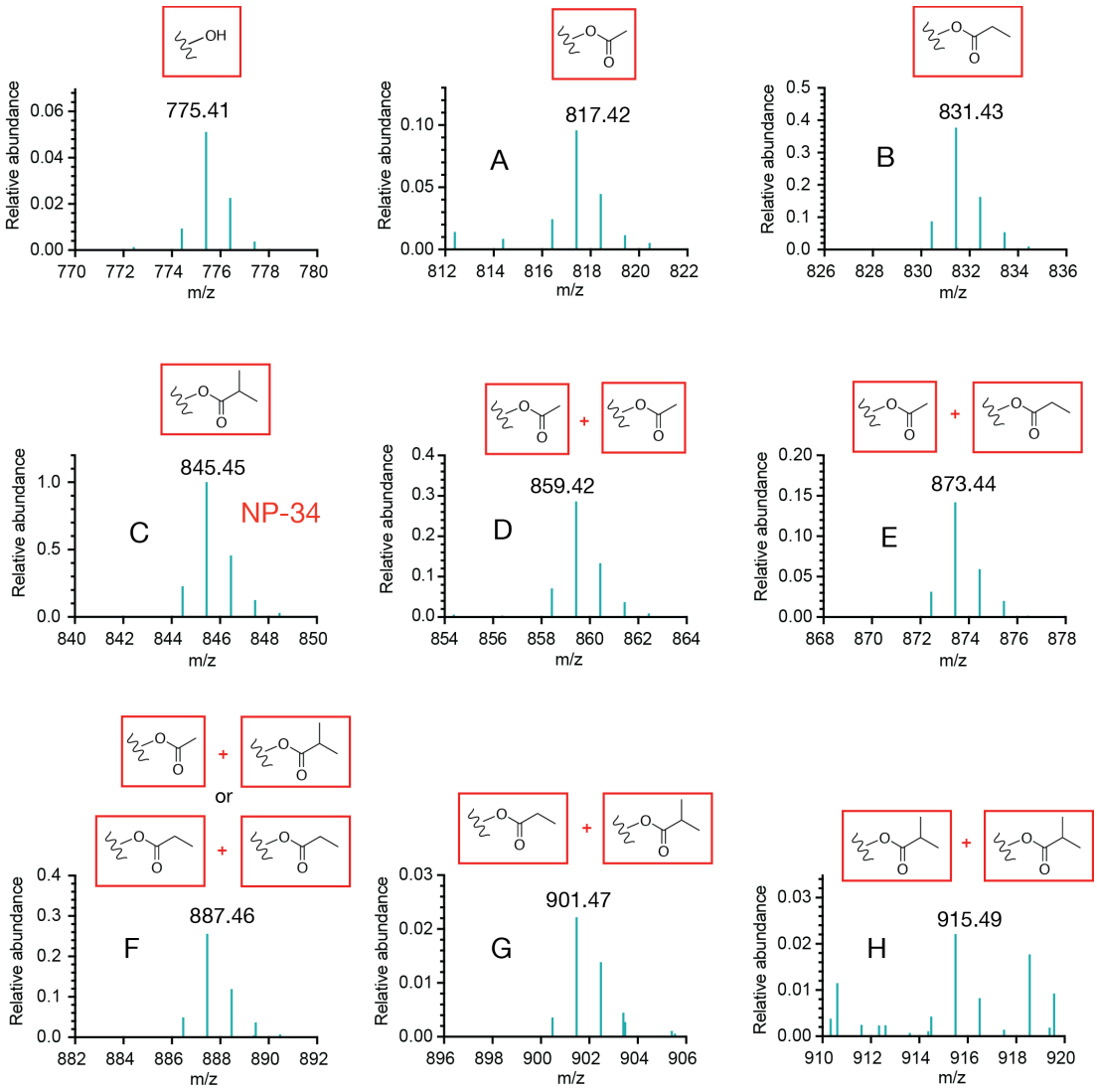


Figure 3.6. HRMS spectra of analogs produced by the proposed BGC of NP-34 produced in *S. albus* J1074.
(A-H) Observation of the characteristic boron splitting pattern corresponding to each of the predicted analogs by HRMS.

3.2.2 Downstream acyltransferase controls production of NP-34 analogs

A knockout of the downstream AT was generated to test the hypothesis that this AT controls the production of NP-34 analogs. Organic extracts from NP-34 heterologously expressed in *S. albus* J1074 and the

downstream acyltransferase knockout were tested by ^{11}B-NMR and HRMS. As expected, ^{11}B-NMR showed a peak around 10.5 ppm indicating the presence of a coordinated boron as previously seen with NP-34. By HRMS, the acyltransferase knockout exhibits only the precursor of NP-34 at m/z = 775.41 while NP-34 and other analogs are not observed. These results indicate the acyltransferase knockout produced a 100 % unacylated product vs. a 11 % unacylated product in the original NP-34 BGC heterologously expressed in *S. albus*.

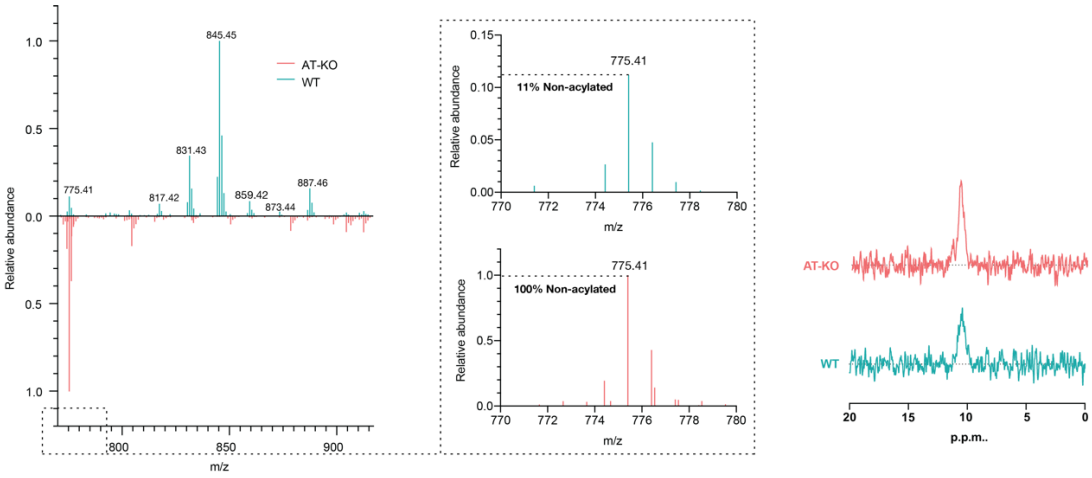


Figure 3.7. ^{11}B-NMR and HRMS of NP-34 BGC with acyltransferase knockout heterologously expressed in *S. albus*.
HRMS of NP-34 BGC in *S. albus* J1074 (WT) and NP-34 BGC with acyltransferase knockout in *S. albus* J1074 (AT-KO). Only the precursor molecule at m/z = 775.41 is observed in AT-KO vs the analogs at m/z = 845.45, 831.43, 859.42, 887.44, 873.44, 817.42, 901.47, and 915.49 in the WT. 100 % of products are unacylated in AT-KO vs. 11 % in WT.

3.2.3 Thioesterase knockout inhibits dimerization of NP-34

Since thioesterases are generally thought to cyclize and cut the growing polyketide chain off the ACP, it was expected a knockout would result in the loss of the full macrolide NP-34 and analogs. Based on ^{11}B-NMR and HRMS, no NP-34 or boronated molecules were observed in the extracts of the thioesterase knockout in the NP-34 BAC expressed in *S. albus* J1074 (Figure 3.8). Additional HRMS analysis and isolation would be beneficial to determine if the monomer can be detected since the growing polyketide chain should be intact until this step. Whether the monomer is able to stay intact is unknown since it would need to be released from the phosphopantetheine arm of the ACP. Without the release of the polyketide chain, the PKS may be stalled and therefore can not generate more polyketide monomers.

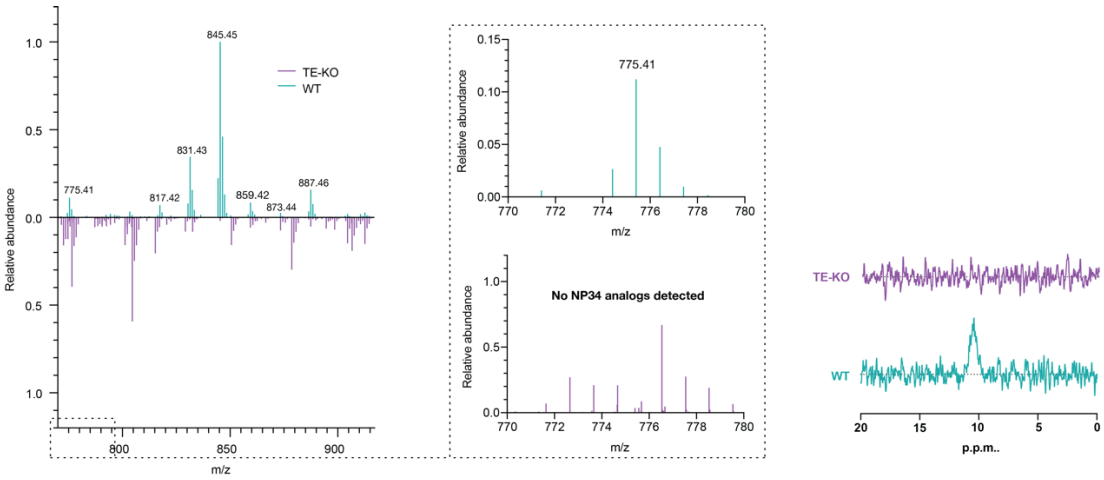


Figure 3.8. HRMS and ^{11}B-NMR of the thioesterase knockout in the NP-34 BAC expressed in *S. albus*.
Comparison of the thioesterase knockout, TE-KO (purple) and WT (blue) NP-34 BAC in *S. albus* J1074 by HRMS and ^{11}B-NMR. Compared to the known analogs of NP-34 at m/z = 845.45, 831.43, 859.42, 887.44, 873.44, 817.42, 901.47, and 915.49, none are detected in TE-KO by HRMS. No boronated products are observed in TE-KO vs WT in ^{11}B-NMR.

3.2.4 LuxR is required for NP-34 production

LuxR, a known transcription factor and suspected participant in quorum sensing, was knocked out between the thioesterase and acyltransferase. If LuxR is a transcriptional repressor, NP-34 biosynthesis could be upregulated whereas NP-34 biosynthesis could be downregulated if it is a transcriptional activator. In this case, NP-34 biosynthesis was completely absent based on HRMS. This could suggest LuxR is a transcriptional activator and is required for transcription of the BGC of NP-34.

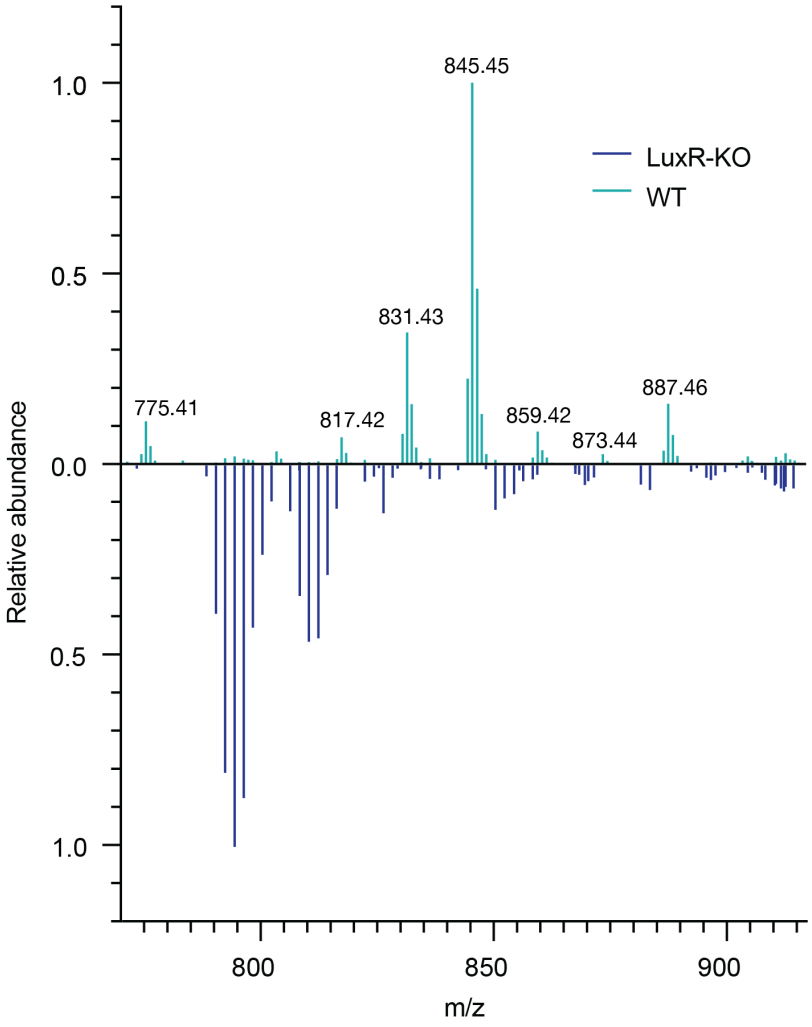


Figure 3.9. HRMS of the LuxR knockout in the NP-34 BAC expressed in *S. albus* J1074.
The LuxR knockout (dark blue) shows no detectable NP-34 analogs as compared to WT (blue).

3.3 Discussion

Generating knockouts in BACs is a valuable tool for understanding the functions of proteins in BGCs that cannot be manipulated in their original organism. By replacing targeted genes with antibiotic selection genes such as AmpR, the screening of knockouts for this study was relatively simple and reliable. This technique was also valuable in screening exconjugates after conjugation of the KO BAC in *S. albus* J1074.

From these studies, it was determined that the TE and LuxR genes in the NP-34 BAC are required for the complete synthesis of NP-34. As mentioned previously, the TE could be responsible for dimerization, cyclization, and cutting the growing polyketide chain off the ACP. Knockout of this gene prevents NP-34 production in agreement with this hypothesis. Knockout of the LuxR gene similarly disrupts the production of NP-34, suggesting that heterologous expression of the NP-34 BAC is somehow regulated by the LuxR genes. This is not unexpected because LuxR genes have been implicated in the regulation of the synthesis of other natural products, including those involved in quorum sensing.[19–21]

These studies also confirmed that the downstream acyltransferase is responsible for the last step of transferring a variety of acyl groups onto either of the two available hydroxyl groups after the polyketide has been dimerized and boronated. Confirmation of the function of the downstream AT opens up a multitude of possibilities in manipulating the macrolide further. Chapter 4

was inspired by these results. If different acyl groups can be incorporated onto the macrolide, there is a possibility an alkyne handle can be incorporated. This opens up the possibility of creating many different analogs of NP-34 through click chemistry. This provides a framework to further develop NP-34 analogs that enhance the specificity of the molecule as an insecticide solely against mosquitos. Trying to chemically change functional groups at this hydroxyl position on the macrolide is very difficult due to the presence of 2 methyl groups causing steric hinderance. Because of this, having the opportunity to perform chemistry at this position biochemically is advantageous.

Knockout of the downstream AT also provides an opportunity to selectively produce the aplasmomycin precursor molecule. This provides a valuable starting material that could be selectively derivatized to produce a variety of NP-34 analogs. As mentioned earlier in this chapter, recombinant expression of the catalytic domain of the downstream AT could be used to produce specific NP-34 analogs based on biocatalysis of acyl substrates and the aplasmomycin precursor molecule. This study provides a framework to selectively produce an aplasmomycin precursor using heterologous expression of the NP-34 BAC with a knockout of the downstream AT.

3.5 References

(1) Yang, B. Three Scientists Who Invented Gene Knockout Technology Won Lasker Basic Medical Research Award. *Discov. Med.* **2009**.

(2) Capecchi, M. R. Altering the Genome by Homologous Recombination. *Science* **1989**, *244* (4910), 1288–1292. https://doi.org/10.1126/science.2660260.

(3) Li, X.; Heyer, W.-D. Homologous Recombination in DNA Repair and DNA Damage Tolerance. *Cell Res.* **2008**, *18* (1), 99–113. https://doi.org/10.1038/cr.2008.1.

(4) Rocha, E. P. C.; Cornet, E.; Michel, B. Comparative and Evolutionary Analysis of the Bacterial Homologous Recombination Systems. *PLOS Genet.* **2005**, *1* (2), e15. https://doi.org/10.1371/journal.pgen.0010015.

(5) Chen, Z.; Yang, H.; Pavletich, N. P. Mechanism of Homologous Recombination from the RecA–ssDNA/dsDNA Structures. *Nature* **2008**, *453* (7194), 489–494. https://doi.org/10.1038/nature06971.

(6) Doudna, J. A.; Charpentier, E. The New Frontier of Genome Engineering with CRISPR-Cas9. *Science* **2014**, *346* (6213), 1258096. https://doi.org/10.1126/science.1258096.

(7) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and Specific Genetic Interference by Double-Stranded RNA in

Caenorhabditis Elegans. *Nature* **1998**, *391* (6669), 806–811. https://doi.org/10.1038/35888.

(8) Joung, J. K.; Sander, J. D. TALENs: A Widely Applicable Technology for Targeted Genome Editing. *Nat. Rev. Mol. Cell Biol.* **2013**, *14* (1), 49–55. https://doi.org/10.1038/nrm3486.

(9) N. Helfrich, E. J.; Piel, J. Biosynthesis of Polyketides by Trans -AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33* (2), 231–316. https://doi.org/10.1039/C5NP00125K.

(10) Khosla, C. Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem.* **2009**, *74* (17), 6416–6420. https://doi.org/10.1021/jo9012089.

(11) Kai Blin, Simon Shaw, Hannah E Augustijn, Zachary L Reitz, Friederike Biermann, Mohammad Alanjary, Artem Fetter, Barbara R Terlouw, William W. Metcalf, Eric JN Helfrich, Gilles P van Wezel, Marnix. H Medema & Tilmann Weber. antiSMASH 7.0: New and Improved Predictions for Detection, Regulation, Chemical Structures, and Visualisation. *Nucleic Acids Res.* **2023**. https://doi.org/doi: 10.1093/nar/gkad344.

(12) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

(13) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858. https://doi.org/10.1038/nprot.2015.053.

(14) Zhou, Y.; Tao, W.; Qi, Z.; Wei, J.; Shi, T.; Kang, Q.; Zheng, J.; Zhao, Y.; Bai, L. Structural and Mechanistic Insights into Chain Release of the Polyene PKS Thioesterase Domain. *ACS Catal.* **2022**, *12* (1), 762–776. https://doi.org/10.1021/acscatal.1c04991.

(15) Seshasayee, A. S. N.; Sivaraman, K.; Luscombe, N. M. An Overview of Prokaryotic Transcription Factors. In *A Handbook of Transcription Factors*; Hughes, T. R., Ed.; Subcellular Biochemistry; Springer Netherlands: Dordrecht, 2011; pp 7–23. https://doi.org/10.1007/978-90-481-9069-0_2.

(16) van Kessel, J. C.; Ulrich, L. E.; Zhulin, I. B.; Bassler, B. L. Analysis of Activator and Repressor Functions Reveals the Requirements for Transcriptional Control by LuxR, the Master Regulator of Quorum Sensing in Vibrio Harveyi. *mBio* **2013**, *4* (4), 10.1128/mbio.00378-13. https://doi.org/10.1128/mbio.00378-13.

(17) Li, Y.; Lin, Z.; Huang, C.; Zhang, Y.; Wang, Z.; Tang, Y.; Chen, T.; Zhao, X. Metabolic Engineering of Escherichia Coli Using CRISPR–Cas9 Meditated Genome Editing. *Metab. Eng.* **2015**, *31*, 13–21. https://doi.org/10.1016/j.ymben.2015.06.006.

(18) Murphy, K. C. λ Recombination and Recombineering. *EcoSal Plus* **2016**, *7* (1), 10.1128/ecosalplus.ESP-0011–2015. https://doi.org/10.1128/ecosalplus.esp-0011-2015.

(19) Barnard, A. M. L.; Bowden, S. D.; Burr, T.; Coulthurst, S. J.; Monson, R. E.; Salmond, G. P. C. Quorum Sensing, Virulence and Secondary

Metabolite Production in Plant Soft-Rotting Bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* **2007**, *362* (1483), 1165–1183. https://doi.org/10.1098/rstb.2007.2042.

(20) Hassan, R.; Shaaban, M. I.; Abdel Bar, F. M.; El-Mahdy, A. M.; Shokralla, S. Quorum Sensing Inhibiting Activity of Streptomyces Coelicoflavus Isolated from Soil. *Front. Microbiol.* **2016**, *7*, 659. https://doi.org/10.3389/fmicb.2016.00659.

(21) Latifi, A.; Winson, M. K.; Foglino, M.; Bycroft, B. W.; Stewart, G. S. A. B.; Lazdunski, A.; Williams, P. Multiple Homologues of LuxR and LuxI Control Expression of Virulence Determinants and Secondary Metabolites through Quorum Sensing in Pseudomonas Aeruginosa PAO1. *Mol. Microbiol.* **1995**, *17* (2), 333–343. https://doi.org/10.1111/j.1365-2958.1995.mmi_17020333.x.

Chapter 4

Exploring the promiscuity of the downstream acyltransferase

4.1 Introduction

Previous studies have shown that acyl-coenzyme A's (acyl-CoA) are substrates used by acyltransferases to transfer an acyl group onto a hydroxyl group in primary and secondary metabolism in bacteria.[1,2,3] These acyl groups can vary giving rise to a variety of different products. As shown in the previous chapter, the downstream acyltransferase is responsible for transferring a variety of acyl groups to create different analogs of NP-34. The goal of this study is to test whether novel acyl groups can be incorporated into the biosynthetic pathway of NP-34 to create new analogs.

4.1.1 Initial in vitro experiments with the downstream acyltransferase

My original approach to altering the acyl group was designed to be in vitro. I tried previously to express the acyltransferase protein recombinantly, but after multiple attempts could not get expression of the acyltransferase. Phyre2, a program to predict secondary structure, predicted the acyltransferase to be a trans-membrane protein.[4] As mentioned in chapter 2, transmembrane proteins are difficult to express and purify and usually require purification conditions that mimic the membrane environment. While the full-length AT could not be purified in this study, future experiments could be optimized as well as trying to express and purify a truncated AT of just the catalytic domain.

4.1.2 Using N-acetyl-cysteamine thioesters in place of coenzyme A

The second approach was to use N-acetyl-cysteamine thioesters (SNACs). SNACs have been used to mimic CoA and are a truncated version of CoA (Figure 4.1). It has been shown that some acyltransferases can use acyl-SNACs in place of acyl-CoAs as their substrate in vitro and in vivo.[5] In vitro experiments include Zhou et al., 2010 using a SNAC to validate a proposed biosynthesis of a whole PKS/NRPS.[6] Heine et al., 2015 used SNACs to study substrate flexibility and stereochemistry of a single PKS/NRPS module.[7] A single enzyme, TE, from a PKS/NRPS domain was used to determine macrolide formation and produce a unnatural analogue.[8] While in vitro studies are useful, it can be difficult to express and purify enzymes outside the cell. In vivo feeding studies can bypass that purification step for enzymes that are difficult to express, such as a trans-membrane acyltransferase. In vivo studies have been successfully done by Liu et al., 1998 in conjunction with ^{13}C labeled SNACs to prove a biosynthetic step.[9] Other studies have used knockout strains and fed in SNACs to determine a biosynthetic step.[10]

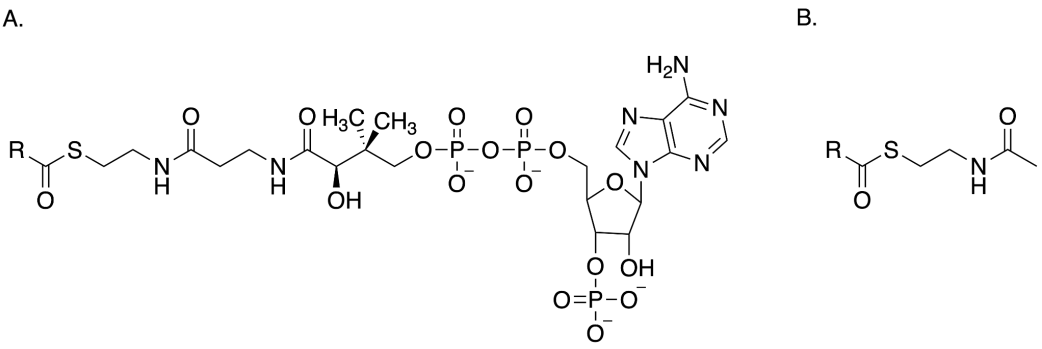


Figure 4.1. Comparison of acyl-CoA versus acyl-SNAC.
A. Structure of acyl-CoA B. Structure of acyl-SNAC, a truncated version of acyl-CoA

In contrast to in vitro studies, a major concern of in vivo studies is the stability of the SNAC in cultures and cellular uptake of the SNAC. If the SNAC degrades or reacts with other substrates before cellular uptake, the reaction cannot take place. Cellular uptake can also be inhibited by the size and functional groups of the SNAC. The reason SNACs are used instead of CoAs is the ability of SNACs to cross cellular membranes while CoAs are too bulky and reactive (Figure 4.1). A major concern of both in vitro and in vivo studies is the correct temperature, media, pH, time of adding SNAC to the experiment, and length of experiment.[5]

4.1.3 Synthesis of SNAK

The acyl group chosen for this study has an alkyne handle. An alkyne handle was chosen due to the flexibility of using click chemistry to “click” or join together an alkyne and an azide. Barry Sharpless at Scripps Research Institute and Morten Meldal at the University of Copenhagen coined the term click chemistry in 2001.[11,12] Along with Carolyn Bertozzi at Stanford

University, all three won the Nobel Prize in Chemistry in 2022 for click chemistry and the development of bioorthogonal reactions. Bertozzi pioneered the way for probing and understanding biological processes without disrupting normal chemical reactions in the cell.[13] For this reason, click chemistry can be used in vivo. If an analog of NP-34 can be made with an alkyne, the possibility of different analogs that can be produced is huge. Without click chemistry, synthesis of these analogs would be quite tedious if at all possible due to the reactivity, stability, and steric hinderance of NP-34 analogs.

In this study I synthesized S-(2-(pent-4-ynamido)ethyl ethanethioate) (SNAK), as done previously.[14] The reaction was adapted from Wang et al., 2017 and is a coupling between N-acetylcysteamine and 4-pentynoic acid in the presence of EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) and DMAP (dimethylaminopyridine) in THF (tetrahydrofuran) (Figure 4.2). The purified SNAK was then used in feeding studies with *S. albus* expressing NP-34 to probe whether the alkyne handle could be incorporated by the acyltransferase as a new analog of NP-34 (Figure 4.3). Since there is a possibility that SNC-034 could be genetically intractable, *S. albus* expressing the NP-34 BGC was used in this study instead of the wildtype strain. Other important factors considered in the study included the amount of SNAK added and at what growth stage. Since the stability and toxicity of SNAK in the media and culture was unknown two different

concentrations of SNAK, 10 mg and 100 mg were added at two different time points, Day 0 and Day 6. Since *S. albus* must be induced for expression of NP-34, Day 6 was chosen, the same day the cultures were induced.

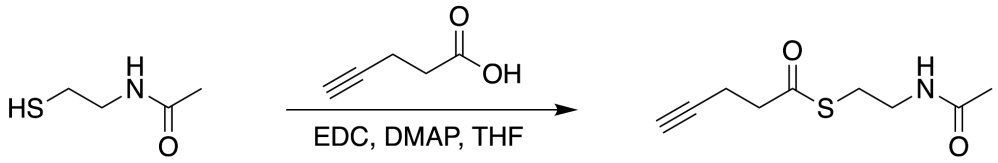


Figure 4.2 Synthesis of SNAK.
Coupling reaction between N-acetylcysteamine and 4-pentynoic acid in the presence of EDC and DMAP in THF to produce SNAK

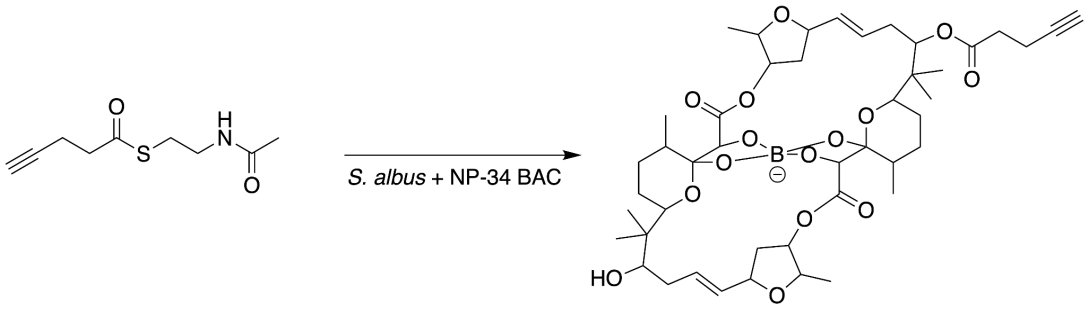


Figure 4.3 Feeding study with SNAK.
Proposed structure of an NP-34 analog containing an alkyne handle supplied by the reaction with SNAK and the NP-34 precursor (m/z = 775.41) from the downstream acyltransferase in the NP-34 BGC (m/z = 855.43).

4.2 Results

4.2.1 Synthesis of the acyl-CoA mimic SNAK with an alkyne handle

The ^{1}H-NMR spectra of SNAK is displayed in Figure 4.4 along with the corresponding chemical shifts. Table # compares chemical shifts to previously published data from Wang et al.[14] Compared to literature values, all shifts are as expected. HRMS confirmed the presence of SNAK with a sodium adduct in positive mode $[M+Na]^{+}$ = 222.05.

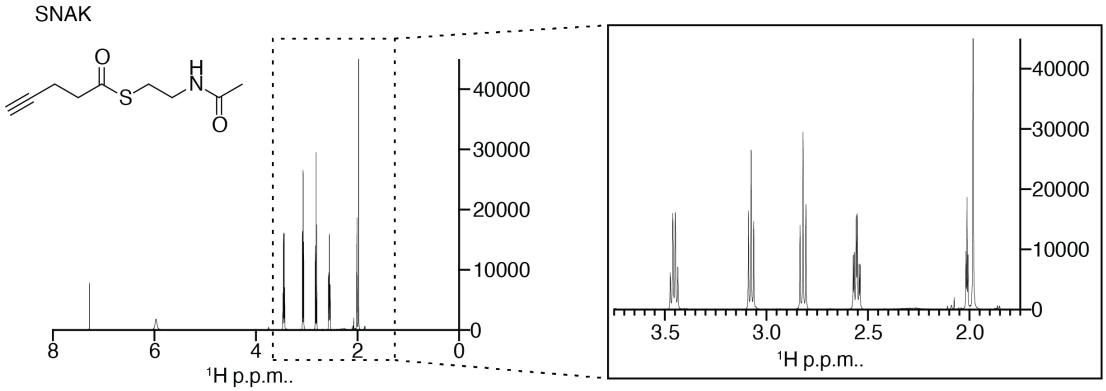


Figure 4.4 ^{1}H-NMR of SNAK in $CDCl_3$.

Chemical Shift	Literature	Experimental	Δ
(s, 1H, NH)	5.81	5.95	0.14
(q, J = 6.1 Hz, 2H, HNCH2CH2S)	3.44	3.43	0.01
(t, J = 6.4 Hz, 2H, HNCH2CH2S)	3.06	3.05	0.01
(t, J =7.2 Hz, 2H, CH2CH2CCH)	2.81	2.8	0.01
(td, J = 7.2, 2.6 Hz, 2H, CH2CH2CCH)	2.54	2.53	0.01
(t, J = 2.6 Hz, 1H, CH2CH2CCH)	1.99	1.99	0
(s, 3H, Acetyl)	1.96	1.96	0

Table 2. Chemical shift of ^{1}H-NMR of SNAK in $CDCl_3$
Comparison of ^{1}H-NMR shifts of SNAK in $CDCl_3$ compared to previously published data.[14]

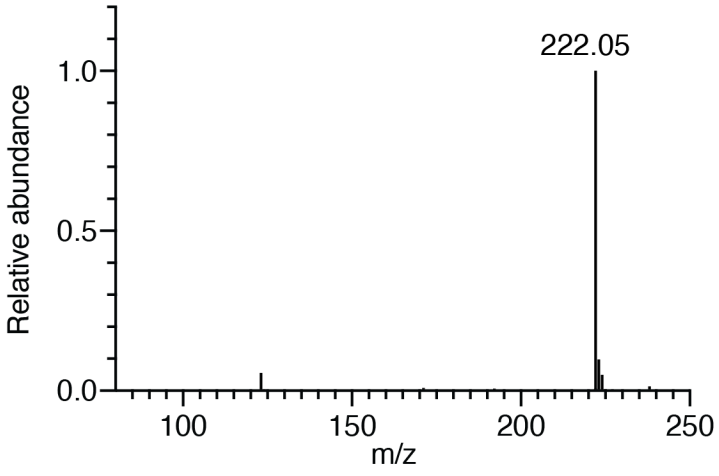


Figure 4.5 HRMS spectra of SNAK in MeOH in positive mode.
$[M+Na]^{+}$ = 222.05

4.2.2 Feeding studies with SNAK

Feeding studies with SNAK did not yield the intended product by HRMS. Organic crude extracts with SNAK yielded two unknown products at m/z = 271.03 and m/z = 863.06 in negative mode. M/z = 271.03 was significantly more concentrated than any other product in the extract. It is unknown what this product is since our SNAK should be a smaller mass (Figure 4.6). In terms of growth, 100 mg of SNAK added at Day 0 led to the culture growing slower than when only 10 mg was added. This could indicate, SNAK is slightly toxic at higher concentrations. Although growth was inhibited, organic extracts still contained NP-34 analogs at low concentrations (Figure 4.7).

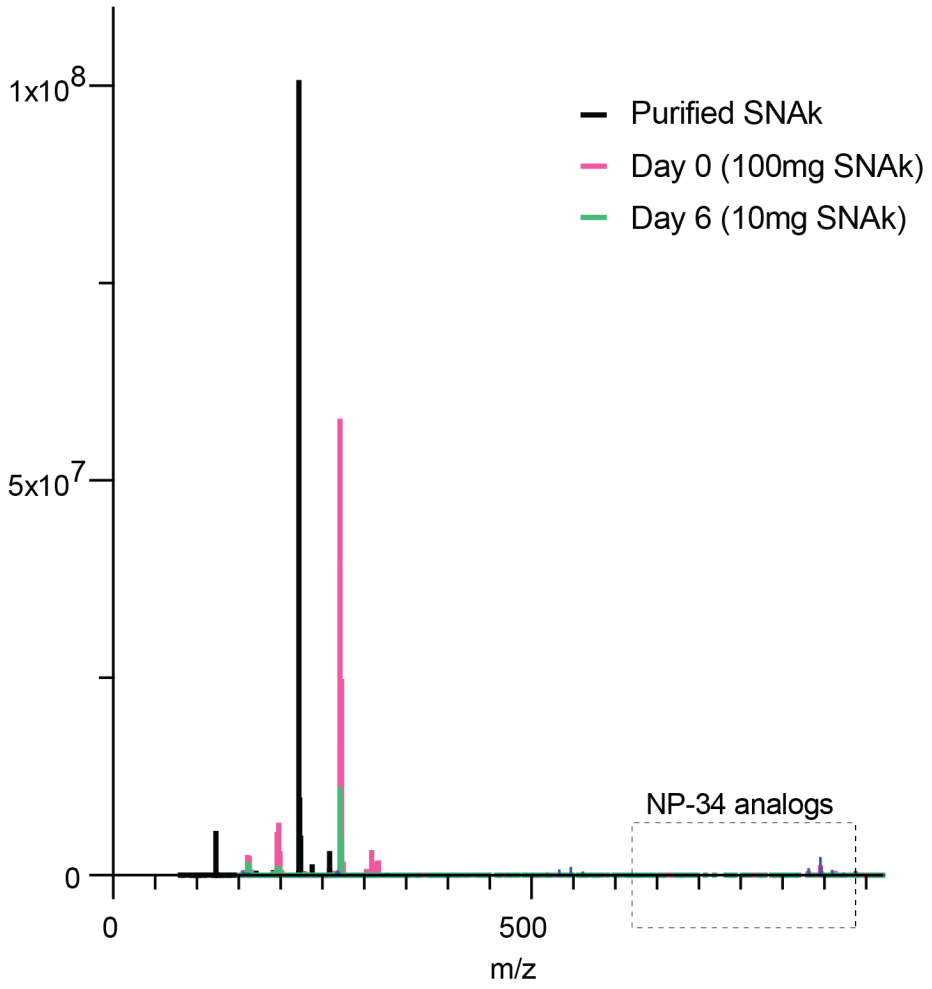


Figure 4.6. HRMS spectra of organic extracts of SNAK feeding study in *S. albus* expressing NP-34 BAC in negative mode.
Purified SNAK (black) vs. 100 mg SNAK at Day 0 (pink) and 10 mg of SNAK at Day 6 (green). NP-34 analogs were produced at a significantly lower concentration than an unknown product at m/z = 271.03

After zooming into the spectra (Figure 4.7), an unknown mass at m/z = 863.06 is observed in the feeding study, but not in the control. This unknown product could be a boronated molecule, but the expected structure is not easily predicted. Isolation and characterization of both unknown products should be done in order to fully characterize each molecule. The unknown

molecule at m/z = 271.03 could be the result of SNAK reacting with another molecule to produce a new product. In result, this could be producing a new analog at m/z = 863.06.

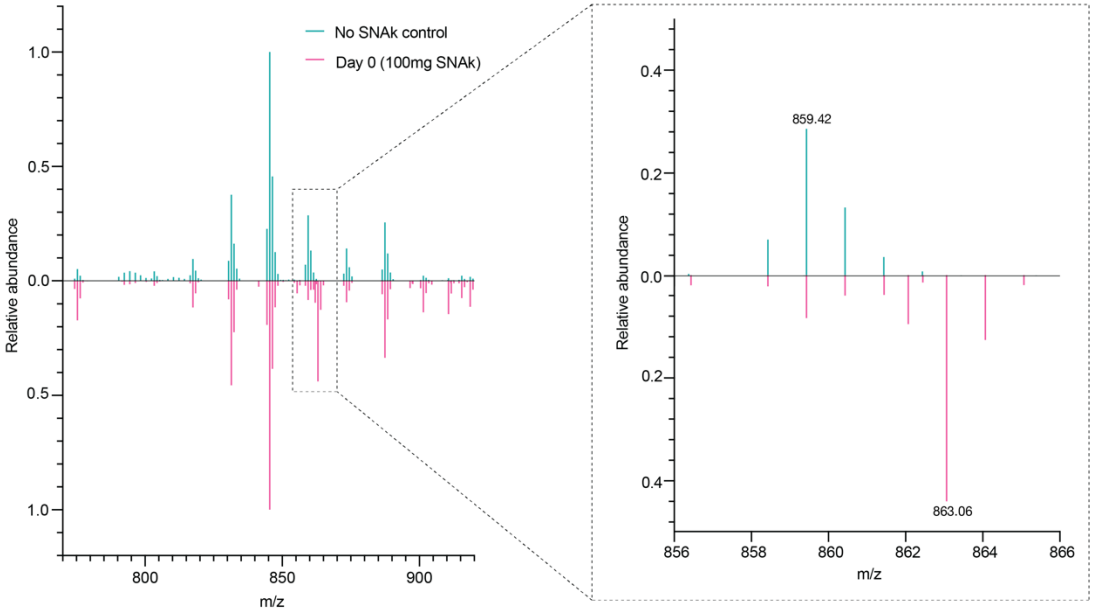


Figure 4.7. Zoomed in HRMS spectra of organic extracts of SNAK feeding study in *S. albus* expressing NP-34 BAC in negative mode. Observed analogs of NP-34 are consistent in both the control (blue) and SNAK feeding study (pink) except for a new unknown mass at m/z = 863.06.

4.3 Discussion

Although the original experimental design did not yield our expected product, there is still potential for this method to produce new analogs of NP-34. Future studies should include isolation and characterization of the unknown products produced. SNAK with the alkyne handle could be reacting in the culture before incorporation by the acyltransferase. If SNAK is too reactive, in vitro studies can also be explored. This could include purifying the acyltransferase protein in favorable conditions for trans-membrane proteins such as a membrane mimicking environment. The unacylated NP-34 analog could then be introduced in the presence of a SNAC to create a new acylated analog.

Other factors to be considered for this project are the feeding schedule and the amount of SNAC added/how often. SNACs may degrade easily in media so adding SNAC once a day could help gradually incorporate this product. Adding 100 mg of SNAK at Day 0 led to delayed growth and decrease in overall cell growth vs. all other conditions. Determining what day and how much SNAK should be added for optimal growth should be explored. Since induction is required for expression of the NP-34 BGC, induction times and time after induction before organic extraction may also influence incorporation of this molecule. Since synthesis of SNAK was a simple one step coupling reaction, other SNAC analogs can be easily made and tested.

Optimization of these factors could greatly improve this experiment and future studies with other SNACs.

References

(1) N. Helfrich, E. J.; Piel, J. Biosynthesis of Polyketides by Trans -AT Polyketide Synthases. *Natural Product Reports* **2016**, *33* (2), 231–316. https://doi.org/10.1039/C5NP00125K.

(2) Khosla, C. Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem.* **2009**, *74* (17), 6416–6420. https://doi.org/10.1021/jo9012089.

(3) Röttig, A.; Steinbüchel, A. Acyltransferases in Bacteria. *Microbiology and Molecular Biology Reviews* **2013**, *77* (2), 277–321. https://doi.org/10.1128/mmbr.00010-13.

(4) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat Protoc* **2015**, *10* (6), 845–858. https://doi.org/10.1038/nprot.2015.053.

(5) Franke, J.; Hertweck, C. Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges. *Cell Chemical Biology* **2016**, *23* (10), 1179–1192. https://doi.org/10.1016/j.chembiol.2016.08.014.

(6) Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W.-H.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y. Enzymatic Synthesis of Resorcylic Acid Lactones by Cooperation of Fungal Iterative Polyketide Synthases Involved in Hypothemycin Biosynthesis. *J. Am. Chem. Soc.* **2010**, *132* (13), 4530–4531. https://doi.org/10.1021/ja100060k.

(7) Heine, D.; Sundaram, S.; Bretschneider, T.; Hertweck, C. Twofold Polyketide Branching by a Stereoselective Enzymatic Michael Addition. *Chem. Commun.* **2015**, *51* (48), 9872–9875. https://doi.org/10.1039/C5CC03085D.

(8) Zhou, Y.; Prediger, P.; Dias, L. C.; Murphy, A. C.; Leadlay, P. F. Macrodiolide Formation by the Thioesterase of a Modular Polyketide Synthase. *Angewandte Chemie International Edition* **2015**, *54* (17), 5232–5235. https://doi.org/10.1002/anie.201500401.

(9) Liu, Y.; Li, Z.; Vederas, J. C. Biosynthetic Incorporation of Advanced Precursors into Dehydrocurvularin, a Polyketide Phytotoxin from Alternaria Cinerariae. *Tetrahedron* **1998**, *54* (52), 15937–15958. https://doi.org/10.1016/S0040-4020(98)01003-5.

(10) Mo, S.; Kim, D. H.; Lee, J. H.; Park, J. W.; Basnet, D. B.; Ban, Y. H.; Yoo, Y. J.; Chen, S.; Park, S. R.; Choi, E. A.; Kim, E.; Jin, Y.-Y.; Lee, S.-K.; Park, J. Y.; Liu, Y.; Lee, M. O.; Lee, K. S.; Kim, S. J.; Kim, D.; Park, B. C.; Lee, S.; Kwon, H. J.; Suh, J.-W.; Moore, B. S.; Lim, S.-K.; Yoon, Y. J. Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues. *J. Am. Chem. Soc.* **2011**, *133* (4), 976–985. https://doi.org/10.1021/ja108399b.

(11) Castelvecchi, D.; Ledford, H. Chemists Who Invented Revolutionary 'Click' Reactions Win Nobel. *Nature* **2022**, *610* (7931), 242–243. https://doi.org/10.1038/d41586-022-03087-8.

(12) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition* **2001**, *40* (11), 2004–2021. https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

(13) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angewandte Chemie International Edition* **2009**, *48* (38), 6974–6998. https://doi.org/10.1002/anie.200900942.

(14) Wang, Y.; Lazor, K. M.; DeMeester, K. E.; Liang, H.; Heiss, T. K.; Grimes, C. L. Postsynthetic Modification of Bacterial Peptidoglycan Using Bioorthogonal *N* -Acetylcysteamine Analogs and Peptidoglycan *O* -Acetyltransferase B. *J. Am. Chem. Soc.* **2017**, *139* (39), 13596–13599. https://doi.org/10.1021/jacs.7b06820.

Materials and methods

Isolation of SNC-034 from marine bacteria

The MacMillan lab isolated SNC-034 from marine sediment, collected from Vava'u, Tonga, coordinates 18° 39'04" S, 173° 59'21", 5 m deep. The sediment was dried and stamped with gauze onto an agar plate containing acidic media. Bacterial colonies from the plate were re-streaked onto new plates until individual bacterial strains were obtained. 16S rRNA analysis of SNC-034 identified 99.9 % identity to *Streptomyces malachitospinus*.

Fermentation of *Streptomyces* strains

Unless otherwise noted, this is the standard method for growing Streptomyces strains. 50 mL of A1Bfe+C media is inoculated with Streptomyces from either a glycerol stock or spores from an agar plate that was streaked from a glycerol stock. Cultures were allowed to grow for 7 days and either upscaled to 1 L or isolated by liquid-liquid extraction. Cultures that were upscaled were inoculated with 5 mL of the 50 mL culture per liter of A1Bfe+C. Cultures were allowed to grow for 7 days and isolated by XAD7 extraction.

Extracting organic molecules from XAD7 Resin

To extract organic molecules from 1 L liquid cultures, activated XAD7 resin was added directly into the flask and allowed to shake at 250 RPM and 25-30°C for 2 hrs. The resin was collected by filtering cultures with cheesecloth. The cheesecloth with the resin was added to a clean flask and acetone was added to completely cover the cheesecloth and resin. Flasks were covered and allowed to shake at the same speed and temp overnight. The resin was filtered off with Whatman paper (10314751) and the solution was collected in round bottoms. The acetone was removed by a rotovap and extracts were brought up in approximately 10 mL methanol by sonication. The sample was transferred to a scint vial and dried down in a speed vacuum.

Liquid-liquid extraction of organics

Liquid-liquid extraction was used for 50 mL cultures. A 50 mL culture was added to a sepratory funnel with 50 mL of ethyl acetate. The solutions were mixed together and the aqueous and organic layers were allowed to separate. The organic fraction was collected in a round bottom and put aside. The process was repeated two more times with the aqueous fraction. The organic fractions were combined and dried down in a rotovap.

Preparation of electrocompetent cells in *E. coli*

This protocol was adapted from the Ayre lab.[1] All *E. coli* competent cells were prepared by the following method. A colony grown on LB agar was selected and grown overnight in LB at 37°C and 250 RPM. 250 mL LB was inoculated with 2.5 mL of overnight culture and grown to an OD_{600} of 0.4-0.5 at 37°C and 250 RPM. Once the OD was reached, the flask was cooled on ice for 15 mins. The contents were transferred to 6, 50 mL sterile falcon tubes and spun down at 4°C and 5,000 RPM in a table top centrifuge. The supernatant was discarded and the cells were resuspended in 250 mL ice cold 10 % glycerol. The process was repeated 3 times with decreasing amounts of glycerol: 125 mL, 12.5 mL, and finally 0.75 mL. The final cell mixture was transferred to a repeater pipet. 50 uL aliquots were dispensed into sterile microcentrifuge tubes on ice. The cells were frozen with liquid nitrogen and stored at -80°C until use.

Transformation of plasmids or BACs into *E. coli* by electroporation

This protocol was adapted from the Ayre lab.[1] Competent cells were removed from the -80°C freezer and thawed on ice for approximately 10 mins. Meanwhile electroporation cuvettes were removed from packaging and cooled on ice. 1-5 uL of DNA was pipetted into 50 uL of electrocompetent cells. The mixture was transferred to a chilled electroporation cuvette and electroporated at 1,800 V. 950 uL of SOC was added to the cuvette and the

mixture was transferred to a microcentrifuge tube. The cells were incubated at 37°C and 250 RPM for 1 hr. 50 uL and 200 uL aliquots were plated on selection plates. The remainder of the cells were spun down and resuspended in 200 uL and plated on selection plates. Plates were incubated at 37°C overnight.

Conjugation of NP-34 gene cluster in Streptomyces

The conjugation protocol was adapted from Practical Streptomyces Genetics.[2]

Preparation of *E. coli* ET12567/puz8002 containing NP-34 BAC

The NP-34 BAC was grown overnight at 37°C in LB and 50 ug/mL apramycin. The BAC was extracted using Zymo ZR BAC DNA Miniprep Kit (D4048) according to the manufacturers protocol. The conjugation strain and helper plasmid *E. coli* ET12567/puz8002 was grown on LB agar with 25 ug/mL chloramphenicol and 50 ug/mL kanamycin. ET12567/puz8002 competent cells were prepared as described previously. The BAC was transformed into ET12567/puz8002 by electroporation as described previously. The BAC was selected for on LB agar plates containing 50 ug/mL apramycin, 50 ug/mL kanamycin, and 25 ug/mL chloramphenicol.

A single colony was grown overnight in 5 mL LB containing 50 ug/mL apramycin, 50 ug/mL kanamycin, and 25 ug/mL chloramphenicol. 1 mL of overnight culture was added to 50 mL LB and previously mentioned

antibiotics until the OD_{600} reached 0.4-0.6. The cells were washed twice in 50 mL LB to removed antibiotics and brought up to a final volume of 5 mL.

Preparation of *Streptomyces* for conjugation

Either streptomyces spores or mycelial was used for conjugation. To prepare streptomyces spores, a liquid culture of streptomyces was streaked onto an MS agar plate. Spores were allowed to form and harvested after 14 days. Cells were harvested by adding 2 mL of 20 % glycerol to the plate. Using a cell scraper, the cells were removed and the cell suspension was collected in a cryovial. Cells were stored at -80 °C until use. To prep spores for conjugation, various concentrations of spores were added to 500 uL of 2 x YT broth and heat shocked at 50°C for 10 mins.

Mycelial was prepared by harvesting 3-4 day old cultures on MS agar plates. The cells were removed from the plate by scraping the cells in the presence of 2 mL of 20 % glycerol. The cell suspension was vortexed and various concentrations of cells were aliquoted into centrifuge tubes.

Conjugation of NP-34 BAC into *S. albus*

500 uL each of prepared *E. coli* and streptomyces cells were mixed and vortexed together. The cells were spun down for 1 min and most of the supernatant was removed. The cells were resuspended with remaining supernatant and plated on MS agar + 10 mM $MgCl_2$. The plates were incubated at 30°C overnight. 500 uL containing 1 mg of apramycin and 500 uL containing 0.5 mg of nalidixic acid was added to each plate and allowed to

dry out in a laminar air flow hood. Once dry, the plates were incubated at 30°C for approximately 4-7 days. Cells were re-streaked onto A1bFe+C plates containing 50 ug/mL apramycin and 25 ug/mL nalidixic acid. Note, conjugations with the NP-34 BAC containing a knockout with the AmpR gene were selected with 100 ug/mL ampicillin along with 25 ug/mL nalidixic acid.

Testing exconjugates

Streptomyces colonies growing on selection plates after conjugation were re-streaked individually on separate plates. One colony from each plate was then added to 50 mL A1BFe+C liquid cultures and grown between 25-30°C for 7 days in the presence of 50 ug/mL apramycin and 25 ug/mL nalidixic acid. 1 mL was removed from the culture and genomic DNA was extracted using Promega Wizard Genomic DNA Purification Kit (A1120) according to the manufacturers protocol. 5 mL of the same culture was added to 1 L of A1BFe+C containing 50 ug/mL apramycin and 25 ug/mL nalidixic acid. Cultures were grown for 7 days and organic molecules were extracted according to the organic molecules extraction procedure with XAD7 resin or liquid-liquid extraction protocol.

Confirmation of NP-34 BAC in Streptomyces

To determine if the NP-34 gene cluster was successfully integrated into the Streptomyces genome, PCR, Sanger sequencing, ^{11}B-NMR, and HRMS was utilized.

PCR and Sanger Sequencing

To test for the successful integration of the NP-34 gene cluster, primers were designed (Benchling) for each gene in the core of the cluster. These genes include Sfp, PksE, Trans AT, KS1, KS5, TE, OX, LuxR, and AT. PCR was performed using EconoTaq 2x master mix according the manufacturers protocol. 5 uL of PCR product was run on a 1 % agarose gel at 80 V for 2 hours. If a band was detected at the correct size, the remaining 20 uL PCR reaction was cleaned-up using the NucleoSpin Gel and PCR Clean-up kit (740609.250). Purified samples were sent for Sanger sequencing (Genewiz). Sequencing results were aligned to the SNC-034 genome to confirm the correct sequence in Benchling using Clustal Omega.

^{11}B-NMR

To prep samples for B-NMR, crude extract was sonicated with 600 uL of deuterated chloroform or methanol and transferred to an NMR tube. The sample was run on a 500 MHz NMR using the pulse sequence outlined in Macho et al.[3]

High resolution mass spectroscopy

All mass spectroscopy was run on an Orbitrap. All crude extracts were prepped to a 1 mg/mL concentration in methanol. Samples were spun down in a microcentrifuge and filtered with 0.22 uM filter to remove any insoluble debris. Samples were directly injected into the Orbitrap at 50 uL/min and

collected as a continuous scan. Results were captured and analyzed using Xcalibur software.

Knockouts by homologous recombination

Generation of AmpR drug casettes

Drug cassettes containing the ampicillin resistance gene (AmpR) were used to replace the targeted genes TE, LuxR, and AT. The protocol was adapted from Sawitzke et al.[4] To build the AmpR drug cassette, 70-bp primers were designed with 50-bp homologous to the upstream of the target sequence and 20- bp homologous to the beginning of the AmpR gene from IDT. This process was repeated for the reverse primer downstream of the target. The drug cassettes were generated by PCR using 2x master mix Q5 High Fidelity DNA polymerase from NEB (M0492S) according to the manufacturers protocol. AmpR was amplified from a pET22b based plasmid gifted from the Partch lab.

Homologous recombination

This protocol was adapted from Li et al.[5] The NP-34 BAC was initially transformed into NEB Turbo Competent *E. coli* cells (C2984I) according to the manufacturers protocol. Colonies containing the BAC were selected on LB containing 50 ug/mL apramycin. From a single colony, electrocompetent cells were produced as previously described. pREDCas9 was purchased from addgene (71541).[5] pREDCas9 plasmid was transformed as previously

described except the outgrowth temperature was at 32°C instead of 37°C since pREDCas9 is temperature sensitive. Cells were selected for on LB plates containing 50 ug/mL apramycin and 50 ug/mL spectinomycin.

A single colony was grown in 5 mL LB overnight containing 50 ug/mL apramycin and 50 ug/mL spectinomycin at 32°C. 50 mL of LB containing 50 ug/mL apramycin and 50 ug/mL spectinomycin was inoculated with 1 mL of overnight culture. The culture was grown at 32°C until the OD_{600} reached 0.5. The culture was induced by raising the temperature to 42°C and shaken for 15 mins. The culture was removed the incubator and chilled on ice for 10 mins. The cells were washed with 50 mL cold sterile ddH_2O twice and brought up in a final volume of 2.5 mL. 50 uL aliquots were prepared for electroporation.

100 ng of the drug cassette was electroporated (1,800 V in 0.1 cm cuvette) into 50 uL of prepared cells. Cells were allowed to grow in 3 mL of LB for 3 hrs. Cells were selected for on LB plates containing 100 ug/mL ampicillin. Colonies containing the knockout were verified through PCR using EconoTaq (30033-1) according to the manufacturers protocol and Sanger sequencing through Genewiz. Once the knockout was verified, the plasmid was purified by Zymo ZR BAC DNA Miniprep Kit (D4048) according to the manufacturers protocol. The BAC was conjugated into *S. albus* J1074 as described previously.

Synthesis of S-(2-(pent-4-ynamido)ethyl ethanethioate) (SNAK)

The synthesis of SNAK was adapted from the protocol in Wang et al.[6] 4-Pentynoic acid (615 mg, 1.25 mmol) was mixed with EDC (685 mg, 1.38 mmol) in 85 mL of THF on ice for 5 mins in a round bottomed flask. The flask was removed from ice and N-acetylcysteamine (670 uL, 1.38 mmol) and DMAP (765 mg, 1.25 mmol) were added. The reaction was stirred overnight at room temp. The reaction mixture was filtered to remove white precipitate and the liquid fraction was dried on a rotovap. The product was dissolved in approximately 50 mL of ethyl acetate and added to a separatory funnel. The organic layer was washed with brine 3 times and dried with sodium sulfate. The organic layer was condensed on a rotovap and purified by column chromatography (silica) with 20-80 % ethyl acetate in hexanes. Fractions were tested by thin liquid chromatography (TLC) on silica plates and stained with potassium permanganate. Like fractions were combined and dried in a rotovap. The product was tested by ^{1}H-NMR in deuterated chloroform and HRMS in methanol.

Feeding studies with SNAK in *S. albus* expressing NP-34

Four conditions were tested starting from the following initial set up: 1 mL of culture from a 50 mL flask containing *S. albus* expressing NP-34 was added to 50 mL of A1BFe+C, 50 ug/mL apramycin, and 25 ug/mL nalidixic acid. At day 0, 1 mM (10 mg) and 10 mM (100 mg) of SNAK was added to

two different cultures. At day 6, the same conditions were repeated in two separate cultures, 1 mM (10 mg) and 10 mM (100 mg) of SNAK. All cultures were induced at day 6. Organic molecules were extracted on day 8, as described previously by liquid-liquid extraction. The extracts were tested by HRMS in methanol and ^{11}B-NMR in deuterated chloroform.

Media Recipes

A1BFe+C: (1L recipe)	Amount
Starch	10 g
Yeast Extract	4 g
Peptone	2 g
Calcium carbonate	1 g
$Fe_2(SO_4)_3$ Solution*	5 mL
KBr Solution*	5 mL
75% Seawater/25% dH_2O	1 L
***$Fe_2(SO_4)_3$ Solution:** (dissolve in 1L dH2O)	
$Fe_2(SO_4)_3$ ($6H_2O$)	8 g
***KBr Solution:** (dissolve in 1L dH_2O)	
KBr	20 g

Artificial Seawater Recipe for 20 L:	**Amount**
Mixture 1:	
NaCl	423.80 g
Na_2SO_4	71.00 g
KCl	11.98 g
$NaHCO_3$	3.48 g
KBr	1.73 g
Boric Acid	0.46 g
NaF	0.06 g
Mixture 2:	
$MgCl_2$ $6H_2O$	191.84 g
$CaCl_2$ $6H_2O$	26.88 g
$SrCl_2$	0.44 g

* Add 0.1% of five supplements (pre-made, for recipes see below) to Mixture 2.

Nutrient 1: (dissolve in 1L dH2O)	
$NaNO_3$	46.70 g
Nutrient 2: (dissolve in 1L dH2O)	
$NaH_2PO_4H_2O$	3.09 g
Metals Stock I: (dissolve both in 1L dH_2O)	
$FeCl_3$ $6H_2O$	1.77 g
Na_2EDTA $2H_2O$	3.09 g
Metals Stock II: (dissolve all in 1L dH_2O)	
$ZnSO_4$ $7H_2O$	73 mg
$CoSO_4$ $7H_2O$	9.80 mg
$MnSO_4$ H_2O	540 mg
Na_2MoO_4 $2H_2O$	1.48 mg
Na_2SeO_3	1.73 µg
$NiCl_2$ $6H_2O$	1.49 mg
Na_2EDTA $2H_2O$	2.44 g
Vitamin Stock: (dissolve all in 1L dH_2O); store at +4°C	
Thiamine-HCl	100 mg
Biotin	2 mg
Vitamin B12	1 mg

References

(1) *Ayre Lab Protocols | Department of Biological Sciences*. https://biology.unt.edu/ayre-lab-plant-physiology-biochemistry-and-signaling/ayre-lab-protocols (accessed 2023-11-20).

(2) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; Crowes: Norwich, England, 2000.

(3) Macho, J. M.; Blue, R. M.; Lee, H.-W.; MacMillan, J. B. Boron NMR as a Method to Screen Natural Product Libraries for B-Containing Compounds. *Org. Lett.* **2022**, *24* (17), 3161–3166. https://doi.org/10.1021/acs.orglett.2c00885.

(4) Sawitzke, J. A.; Thomason, L. C.; Bubunenko, M.; Li, X.; Costantino, N.; Court, D. L. Recombineering. In *Methods in Enzymology*; Elsevier, 2013; Vol. 533, pp 79–102. https://doi.org/10.1016/B978-0-12-420067-8.00007-6.

(5) Li, Y.; Lin, Z.; Huang, C.; Zhang, Y.; Wang, Z.; Tang, Y.; Chen, T.; Zhao, X. Metabolic Engineering of Escherichia Coli Using CRISPR–Cas9 Meditated Genome Editing. *Metabolic Engineering* **2015**, *31*, 13–21. https://doi.org/10.1016/j.ymben.2015.06.006.

(6) Wang, Y.; Lazor, K. M.; DeMeester, K. E.; Liang, H.; Heiss, T. K.; Grimes, C. L. Postsynthetic Modification of Bacterial Peptidoglycan Using Bioorthogonal *N* -Acetylcysteamine Analogs and Peptidoglycan *O* -

Acetyltransferase B. *J. Am. Chem. Soc.* **2017**, *139* (39), 13596–13599. https://doi.org/10.1021/jacs.7b06820.