

**UNDERSTANDING AND OVERCOMING INDUCIBLE RIFAMYCIN  
RESISTANCE**

**UNDERSTANDING AND OVERCOMING INDUCIBLE RIFAMYCIN  
RESISTANCE**

By MATTHEW D. SURETTE, B.Sc.

A thesis submitted to the School of Graduate Studies in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy

McMaster University © Copyright Matthew David Surette, April 2023

**DESCRIPTIVE NOTE**

McMaster University DOCTOR OF PHILOSOPHY (2023) Hamilton, Ontario  
(Biochemistry and Biomedical Sciences)

TITLE: Understanding and overcoming inducible rifamycin resistance

AUTHOR: Matthew D. Surette, B.Sc.

SUPERVISOR: Gerard D. Wright, Ph.D.

NUMBER OF PAGES: xiv, 291

## **FOREWORD**

### **LAY ABSTRACT**

Our antibiotic arsenal consists primarily of metabolites produced by soil microbes, which humanity repurposed into life-saving medicines in the 20<sup>th</sup> century. As a direct result of the natural origin of antibiotics, resistant bacteria exist in these same environments, independent of human use. Individual genetic determinants from this reservoir can emerge in pathogenic bacteria without warning and render antibiotics ineffective. The aim of this work was to understand how environmental bacteria resist the rifamycin class of antibiotics. Firstly, I investigated the ability of some bacteria to sense the presence of rifamycins, and in response produce proteins to protect themselves. I discovered that this process requires specific DNA sequences nearby resistance genes. Using this DNA sequence as a guide I cataloged resistance genes in thousands of bacterial genomes and discovered a new mechanism of rifamycin resistance. Lastly, I exploited this rifamycin sensing system to discover new antibiotics from soil microbes.

## ABSTRACT

Antibiotics are one of the most important advances in medical science, but today, antibiotic-resistant bacteria threaten this legacy. We risk losing our ability to treat acute infections, perform invasive surgeries, and exploit immunosuppressive therapies like transplantation and cancer chemotherapy. The antibiotics we use today have ancient roots and have been produced by microbial denizens of the soil for millions of years before we adopted them in the 20<sup>th</sup> century. This history has modern consequences, as strategies to resist these compounds have evolved in concert for millions of years. The result is a vast reservoir of antimicrobial resistance that exists in environmental bacteria, which have the potential to be mobilized into human pathogens and cripple our antibiotic arsenal. Here, I set out to deepen our understanding of the environmental resistome, focusing on the rifamycin antibiotics. These compounds inhibit bacterial RNA polymerase and are frontline agents for treating tuberculosis. Environmental bacteria from the phylum Actinobacteria induce the production of resistance enzymes in response to these compounds. Although mechanistic questions remain, we demonstrate that this induction stems from the inhibition of RNA polymerase by rifamycins. The induction process is known to require a specific DNA motif; here, I identify additional sequences as part of this motif and use this information to map inducible rifamycin resistance across the entire phylum. The most common rifamycin-inducible gene was an uncharacterized family of proteins annotated as DNA helicases. I investigated these proteins and discovered that they bind to RNA polymerase and displace rifamycin antibiotics, a novel mechanism of rifamycin resistance. Lastly, we repurposed this inducible system to develop an assay to screen for novel RNA polymerase inhibitors. From this screen, we identified a rifamycin immune to a common environmental resistance enzyme and a new family of rifamycin antibiotics.

## ACKNOWLEDGEMENTS

I've always believed that science is a social activity. This dissertation reflects the overwhelming support I received from my colleagues, friends, family, and mentors. Some contributions were direct and scientific, and others were indirect, but they were all significant.

Firstly I need to thank Gerry for taking a chance on me and having faith that I'll figure things out eventually. You've probably given me more freedom than I deserved to take this project in my own direction(s). Dr. Lori Burrows and Dr. Marie Elliot, who made up the rest of my supervisory committee were also instrumental. Their unique perspectives, helpful discussion, and feedback helped me mature tremendously as a scientist.

I also need to thank members of my family for their endless patience and support as I spent many nights and weekends in the lab. Haley, my partner, who was with me through all the peaks and valleys of my PhD. My Mom, who's always nurtured my curiosity. My Dad, who's always been my role model for how to be a compassionate and collaborative scientist. Carolyn and Ben, who make up the rest of my home away from home. Willow, my loyal guard dog, who's offered silent support throughout the thesis writing process. I couldn't have done it without you all.

I've also benefitted from exposure to some outstanding colleagues in the Wright Lab. Kalinka, who helped me find my footing early on and taught me so much chemistry, but also about drive and perseverance. (New Delhi) Dave who's a wealth of information about proteins, structural biology, and enzymology but also has an infectious passion for science. Min, who was always willing to entertain big and small questions, and introduced me to synthetic biology. Nick, Andrew (AP), Mike, Allison, Wen, and Linda all imparted some of their wisdom on me as well.

I was blessed with a fantastic cohort during my time in the Wright Lab. It was a pleasure to get to know all of you both inside and outside the lab. Grace, Arthur, Dan, Drew, Vishwas, Venky, Beth, Emily, Allison, Caitlyn, Fei, Li Jun (the unicorn wrangler), Min Wen, Adam, Manpreet, Manoj, Divya, and of course, the Brown Lab Ballers.

## TABLE OF CONTENTS

<b>Chapter 1 Introduction</b>	<b>1</b>
Antibiotics and the Resistome	2
Rifamycin antibiotics	14
Rifamycin resistance	33
Aims of this work	50
<b>Chapter 2 – Induction through the RAE</b>	<b>52</b>
Abstract	54
Introduction	55
Results	58
Discussion	69
Methods	78
References	84
Supplemental material	89
<b>Chapter 3 A genomic survey of inducible rifamycin resistance</b>	<b>91</b>
Abstract	93
Introduction	94
Results	96
Discussion	113
Methods	121
References	125
Supplemental material	128
<b>Chapter 4 HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics</b>	<b>134</b>
Summary	136
Introduction	136
Results	140
Discussion	160
Methods (STAR methods)	169
References	194
Supplementary material	199
<b>Chapter 5 Repurposing inducible rifamycin resistance for the high throughput discovery of RNA polymerase inhibitors</b>	<b>212</b>
Abstract	214
Introduction	215
Results	220
Discussion	238
Methods	243
References	256

Supplementary material	261
<b>Chapter 6 Discussion</b>	<b>271</b>
Summary	272
The role of rifamycins and resistance in natural environments	273
The future of rifamycin antibiotics	275
Concluding remarks	278
References	279

## LIST OF FIGURES

Figure 1 Mechanisms of antibiotic resistance	8
Figure 2 Structure and SAR of rifamycin antibiotics	15
Figure 3 Structures of semi synthetic rifamycins	19
Figure 4 Proposed biosynthesis of rifamycin antibiotics	23
Figure 5 Diversity in natural product rifamycins	25
Figure 6 The transcription cycle	27
Figure 7 The steric occlusion model	30
Figure 8 Rifamycin inhibition of RNA polymerase	31
Figure 9 The rifamycin resistance determining region (RRDR) of RpoB	35
Figure 10 The rifamycin ADP-ribosyltransferase Arr	40
Figure 11 The rifamycin glycosyltransferase Rgt	43
Figure 12 The rifamycin phosphotransferase Rph	45
Figure 13 The rifamycin monooxygenase Rox/Iri	48
Figure 14 The rifamycin-associated element (RAE)	49
Figure 1 The RAEs in <i>S. venezuelae</i> and their transcription start sites	59
Figure 2 A reporter system for rifamycin mediated induction through the RAE	61
Figure 3 Conserved spacing of Extended RAE elements is required for induction	62
Figure 4 Rifamycin resistant RpoB alleles are defective in induction of the RAE	64
Figure 5 Induction of the RAE is specific to antibiotics targeting the rifamycin binding pocket	66
Figure 6 $P_{rox}$ is susceptible to rifampin <i>in vitro</i>	69
Supplemental Figure 1 Purification of streptolydigin	89
Supplemental Figure 2 Fidaxomicin antagonizes induction of the RAE but doesn't display antimicrobial synergy	90
Figure 1 The extended RAE	98
Figure 2 Sequence similarity network of RAE-associated protein families	100
Figure 3 Distribution of the RAE and associated resistance mechanisms in Actinobacteria	105
Figure 4 RAE and rifamycin resistance gene frequency in the Mycobacteraceae	107
Figure 5 Inducible rifamycin resistance in the mycobacteria	108

Figure 6 Prevalence and specific inventory of multi-RAE organisms	110
Figure 7 Rifamycin inactivation is common in soil actinomycetes	112
Supplementary Figure 1 Sequence similarity network of RAE-associated CDS in the 2 and 3 position.	129
Supplementary Figure 2 RAE inventory of <i>Amycolatopsis</i>	130
Supplementary Figure 3 Encoding multiple RAEs is more frequent in specific taxonomic groups within Actinobacteria	132
Supplementary Figure 4 Predicted frequency of rifamycin inactivating enzymes in select Actinobacterial genomes	133
Figure 1 Rifamycin inactivation mechanism are controlled by RAE sequences	141
Figure 2 HelR is not an inactivating enzyme but confers tolerance to rifamycins	146
Figure 3 HelR <sub>Sv</sub> co-purifies with RNA polymerase in <i>S. venezuelae</i> and blocks rifampin activity <i>in vitro</i>	152
Figure 4 Design and activity of a rifamycin photoprobe (RPP) demonstrate that HelR <sub>Sv</sub> displaces RNAP-bound rifamycins	155
Figure 5 Sequence Similarity network of HelD-like proteins	159
Figure 6 Model of HelR-mediated rifamycin resistance	165
Supplementary Figure 1 Induction through the RAE does not require HelR	199
Supplementary Figure 2 Distance between HelR <sub>Ms</sub> PCh loop residues and rifampin	201
Supplementary Figure 3 Rifamycin binding pocket in a State II RNAP:HelR <sub>Ms</sub>	202
Supplementary Figure 4 Supplementary Figure 4. Progress curve for multiple round <i>in vitro</i> transcription using P <sub>PA3</sub>	203
Supplementary Figure 5 RPP labelling controls	204
Supplementary Figure 6 Distribution of HelR clusters in the Sequence Similarity Network of HelD-like proteins	206
Supplementary Figure 7 Rifamycin Photoprobe (RPP) synthetic route and NMR assignment key	207
Figure 1 Inducible rifamycin resistance	219
Figure 2 A high throughput screen for RAE activation	222
Figure 3 Natural product library screen and dereplication	224
Figure 4 Chaxamycin D	226

Figure 5 Purification of an Rph-evasive rifamycin from WAC02994	230
Figure 6 Structural and genetic diversity of rifamycin-like antibiotics	233
Figure 7 WAC09654 rifamycin-like BGC and the predicted product	235
Figure 8 Heterologous expression of WAC09654	237
Supplementary Figure 1 Hypersusceptible <i>S. venezuelae</i> pGUS: $P_{rox}$ improves detection of rifampin significantly	261
Supplementary Figure 2 Cytochalasin D does not inhibit RNAP or induce the RAE	262
Supplementary Figure 3 BGC similarity of WAC09155 and WAC02994 to BGC's with known products	264
Supplementary Figure 4 WAC09654 cluster homology and functional predictions	267
Supplementary Figure 5 WAC09654 cluster capture	268
Supplementary Figure 6 Early rifamycin and WAC09654 biosynthesis	269
Supplementary Figure 7 Putative self-resistance RpoB substitutions in rifamycin-like antibiotic producing organisms	270

## LIST OF TABLES

Table 1 Strains and isolates used in this work	79
Table 1 Miscellaneous RAE-associated genes	101
Supplementary Table 1 Miscellaneous proteins in CDS2 and 3 positions	128
Supplementary Table 2 Actinomycetaeae subdivisions	131
Table 1 Susceptibility testing	144
Key resources table	170
Table S1 Top 10 enriched and depleted RNAP associated proteins	200
Table S2 Clusters with $\geq 1$ member associated with a RAE	205
Table S3 Chemical shifts for Rifamycin B and Rifampin Photo Probe in dms <sub>o</sub> -d <sub>6</sub> , reported in ppm.	208
Table S4 Oligonucleotides used in this study	210
Table 1 Antimicrobial susceptibility testing of chaxamycin D	227
Table 2 Strains and isolates used in this work	243
Supplementary Table 1 Additional hit information	263
Supplementary Table 2 WAC09155 BGC comparison	265
Supplementary Table 3 WAC02994 BGC comparison	266

## LIST OF ABBREVIATIONS

<b>ABC</b>	ATP-binding cassette
<b>ADP</b>	Adenosine diphosphate
<b>ART</b>	ADP ribosyltransferase
<b>ATP</b>	Adenosine triphosphate
<b>BGC</b>	Biosynthetic gene cluster
<b>BLAST</b>	Basic local alignment search tool
<b>BSA</b>	Bovine serum albumin
<b>CDS</b>	Coding DNA sequence
<b>CTP</b>	Cytosine triphosphate
<b>DMF</b>	Dimethylformamide
<b>DMSO</b>	Dimethylsulfoxide
<b>DNA</b>	Deoxyribonucleic acid
<b>DTT</b>	Dithiothreitol
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>FDA</b>	Food and Drug Administration
<b>FDX</b>	Fidaxomicin
<b>FMN</b>	Flavin mononucleotide
<b>GNPS</b>	Global Natural Product Social Molecular Networking Platform
<b>GTP</b>	Guanosine triphosphate
<b>GUS</b>	Beta glucuronidase
<b>HEPES</b>	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
<b>HGT</b>	Horizontal gene transfer
<b>HRMS</b>	High resolution mass spectrometry
<b>HRP</b>	Horseradish peroxidase
<b>IC<sub>50</sub></b>	50% inhibitory concentration
<b>ICE</b>	Integrative and conjugative elements
<b>KAN</b>	Kanamycin
<b>LPS</b>	Lipopolysaccharide
<b>MccJ25</b>	Microcin J25
<b>MeOH</b>	Methanol
<b>MIC</b>	Minimum inhibitory concentration
<b>NAD</b>	Nicotinamide adenine dinucleotide
<b>NCBI</b>	National center for biotechnology information
<b>Ni-NTA</b>	Nickel nitriloacetic acid
<b>NMR</b>	Nuclear magnetic resonance
<b>NTM</b>	Non-tuberculosis mycobacteria
<b>NTP</b>	Nucleotide triphosphate

<b>PCR</b>	Polymerase chain reaction
<b>PDB</b>	Protein databank
<b>PLP</b>	Pyridoxal phosphate
<b>PVDF</b>	Polyvinylidene fluoride
<b>RNA</b>	Ribonucleic acid
<b>RRDR</b>	Rifamycin resistance determining region
<b>RT-qPCR</b>	Reverse-transcriptase quantitative polymerase chain reaction
<b>SAR</b>	Structure activity relationship
<b>SDS-PAGE</b>	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
<b>SOR</b>	Sorangicin
<b>STL</b>	Streptolydigin
<b>TSS</b>	Transcription start site
<b>US</b>	United States
<b>UTP</b>	Uridine triphosphate
<b>UV</b>	Ultraviolet
<b>VRE</b>	Vancomycin resistant Enterococci
<b>WHO</b>	World Health Organization

**DECLARATION OF ACADEMIC ACHIEVEMENT**

I have performed all the research in this body of work except where indicated in the preface of each chapter.

## **CHAPTER 1**

### **Introduction**

## ANTIBIOTICS AND THE RESISTOME

### The discovery of antibiotics

The term chemotherapy, the treatment of diseases with specific individual chemical entities, was coined in the early 20<sup>th</sup> century by the German scientist Paul Ehrlich<sup>1</sup>. Dyes synthesized for the textile industry were widely available to scientists in Germany at the time, and Ehrlich was among the first to observe that certain dyes could preferentially stain specific cell types. He reasoned that if chemicals could discriminate between different human cells, it must be possible to synthesize molecules that could specifically bind and kill disease-causing microbes. A concept he called *chemotheripia specifica*. In Frankfurt, Ehrlich assembled a team resembling the modern pharmaceutical apparatus. Starting from an active but toxic organoarsenic molecule, Atoxyl (arsanillic acid), this team synthesized hundreds of analogs, tested them for safety and efficacy in rabbits, and even ran small human trials. This process yielded what most consider to be the first antibiotic, Salvarsan, for the treatment of syphilis<sup>2,3</sup>. The potential of this approach was immediately apparent.

“Hardly at any time in the history of modern medicine has there existed a more intense excitement and a more absorbing interest among the medical fraternity than at present. One of the greatest scourges of humanity – perhaps the most insidious and cruel of all, since it so often places its victims beyond the pale of human sympathy, to be loathed rather than pitied – is on the point of being eradicated.”

-H. Schweitzer 1910 in Science Magazine, referencing Ehrlich’s Salvarsan<sup>4</sup>

The 1930s and 40s would see an explosion of success molded in the image of Ehrlich’s vision. First, Dr. Gerhard Domagk’s team at Bayer won the Nobel prize in 1939 for discovering and developing the sulfa drug prontosil. Unlike Salvarsan, prontosil

possessed activity against most Gram-positive cocci and could thus be used to treat many infections<sup>5</sup>. Next, a curious observation on an agar plate, the inhibition of *Staphylococcus aureus* by a fungal contaminant, led to one of the most important discoveries of the 20<sup>th</sup> century, penicillin. Alexander Fleming first reported his discovery of penicillin in 1929 but could not isolate the active component in the years following his publication<sup>6,7</sup>. Roughly a decade later, a team led by Howard Florey and Ernst Chain developed a robust methodology for producing and purifying the active component, penicillin F. Initial tests showed extraordinary efficacy in animals and humans<sup>8,9</sup>. Finally, a massive collaborative effort between the United Kingdom and the United States spurred by personnel losses to infections during World War II led to penicillin production on an industrial scale<sup>10</sup>. Fleming, Florey, and Chain won the Nobel prize in 1945 for their discovery of penicillin. In the years following the war, penicillin became widely available, and the age of antibiotics began.

Penicillin was a transformative discovery, but more importantly, it set the stage for exploring microbial metabolites as human medicines. Microbiologists had noted bacterial antagonism for decades, but the tools and techniques to purify and characterize these molecules were in their infancy. This fact and a lack of appreciation for their medical utility relegated them to laboratory and ecological phenomena<sup>11</sup>. In the 1930s, Selman Waksman was the first to survey microbes for antibiotic activity at a large scale and high throughput<sup>12</sup>. Working primarily with filamentous soil bacteria (actinomycetes), his team discovered numerous antibiotics such as neomycin, actinomycin, and streptothricin<sup>13–15</sup>. In 1952 he won the Nobel prize for his discovery of streptomycin, the first antimicrobial therapy for

tuberculosis<sup>16</sup>. The so-called “Waksman platform” was soon imitated worldwide, and the golden age of antibiotic discovery began. Between 1945-1970 researchers discovered almost all the chemical classes of antibiotics currently used in the clinic. Antibiotics found through *de-novo* chemical syntheses such as fluoroquinolones, oxazolidinones, and select anti-tubercular compounds are far less abundant than those derived from microbes<sup>17,18</sup>. For instance, the five most common outpatient antibiotic prescriptions in the US during 2021 were amoxicillin, azithromycin, amoxicillin-clavulanate, doxycycline, and cephalexin; all semi-synthetic derivatives of microbial metabolites<sup>19</sup>.

Infectious diseases were among the leading causes of death in Europe and North America at the turn of the 20<sup>th</sup> century. Tuberculosis, pneumonia/influenza, and diarrheal illness accounted for 30% of deaths in the United States in 1900<sup>20</sup>. The proportion of deaths caused by infectious disease would fall by ~95% between 1900 and 1980<sup>21</sup>. This precipitous drop occurred due to a confluence of factors, including improved nutrition, living standards, food and water safety, improved sanitation, and hygiene practices, widespread immunization, and the introduction of antibiotics. It is hard to quantify the effect of any single intervention, but it is widely agreed that antibiotics have saved countless lives since their introduction in the middle of the 20<sup>th</sup> century. However, the role of antibiotics in medicine goes beyond treating acute infections. Whether they are used prophylactically or following the development of an infection, the presence of antibiotics in the physician’s repertoire has enabled more invasive and routine surgical interventions. Immunosuppressive or cytotoxic therapies, such as those used in transplantation medicine or cancer chemotherapy, are predicated on the colossal shift in the risk-benefit calculus

brought by antimicrobials. In effect, the modern medical system relies on the availability of effective antibiotics.

### **The age of resistance**

At the dawn of the antibiotic era, common human pathogens were susceptible to each new antibiotic in almost all cases. Not all antibiotics could kill all bacteria; for instance, penicillin G had poor activity against Gram-negative pathogens such as *E. coli*. However, penicillin G was effective against essentially all isolates of *Staphylococcus aureus*. This pan-susceptibility is apparent from contemporary accounts and retrospective studies on well-preserved strains from the pre-antibiotic era, such as the Murray collection containing Enterobacteriaceae from 1917 – 1954<sup>22</sup>. Alexander Fleming issued a famous warning during his Nobel prize acceptance speech that penicillin must be prescribed cautiously, as it is easy to generate resistant bacteria in the lab<sup>23</sup>. By the late 1940s, hospitals encountered high levels of untreatable penicillin-resistant *Staphylococci*. At the Hammersmith hospital in London, the rate of penicillin-resistant *Staphylococcus aureus* rose from 14% to 59% in less than two years<sup>24</sup>. The emergence of penicillin-resistant pathogens followed the widespread introduction of penicillin. This pattern would hold for all subsequent antibiotics introduced into the clinic, thankfully almost always at a slower pace than Hammersmith hospital.

Today, the exact prevalence of resistance varies based on socioeconomic factors, geography, and the bacteria and antimicrobial in question, but the impact of resistance is both global and severe in scale. A commission in the United Kingdom predicted that antibiotic resistance could account for 10 million deaths a year by 2050<sup>25</sup>. In line with these

predictions, a recent study estimated that in 2019 alone, almost 5 million deaths were associated with drug-resistant infections and that 1.27 million are directly attributable to antibiotic resistance<sup>26</sup>. Many bacteria possess intrinsic resistance to certain antimicrobials; this trait is common to all species or genera members. In specific cases, intrinsic resistance can restrict treatment options for some organisms, but generally, enough alternatives exist that this isn't a problem. The antibiotic resistance crisis is driven instead by acquired resistance, occurring when a bacterium is no longer susceptible to an antibiotic that is actively used to treat it. The emergence of multi-drug resistant organisms, which are no longer vulnerable to agents within multiple antibiotic classes, is becoming increasingly common. Furthermore, isolates resistant to all available antimicrobials have been cultured from infections but are still very rare<sup>27,28</sup>. These facts have led many to declare that we are in the age of resistance. For instance, this quote by the director of the US Centers for Disease Control and Prevention, Robert Redfield, "Stop referring to a coming post-antibiotic era – it's already here"<sup>29</sup>.

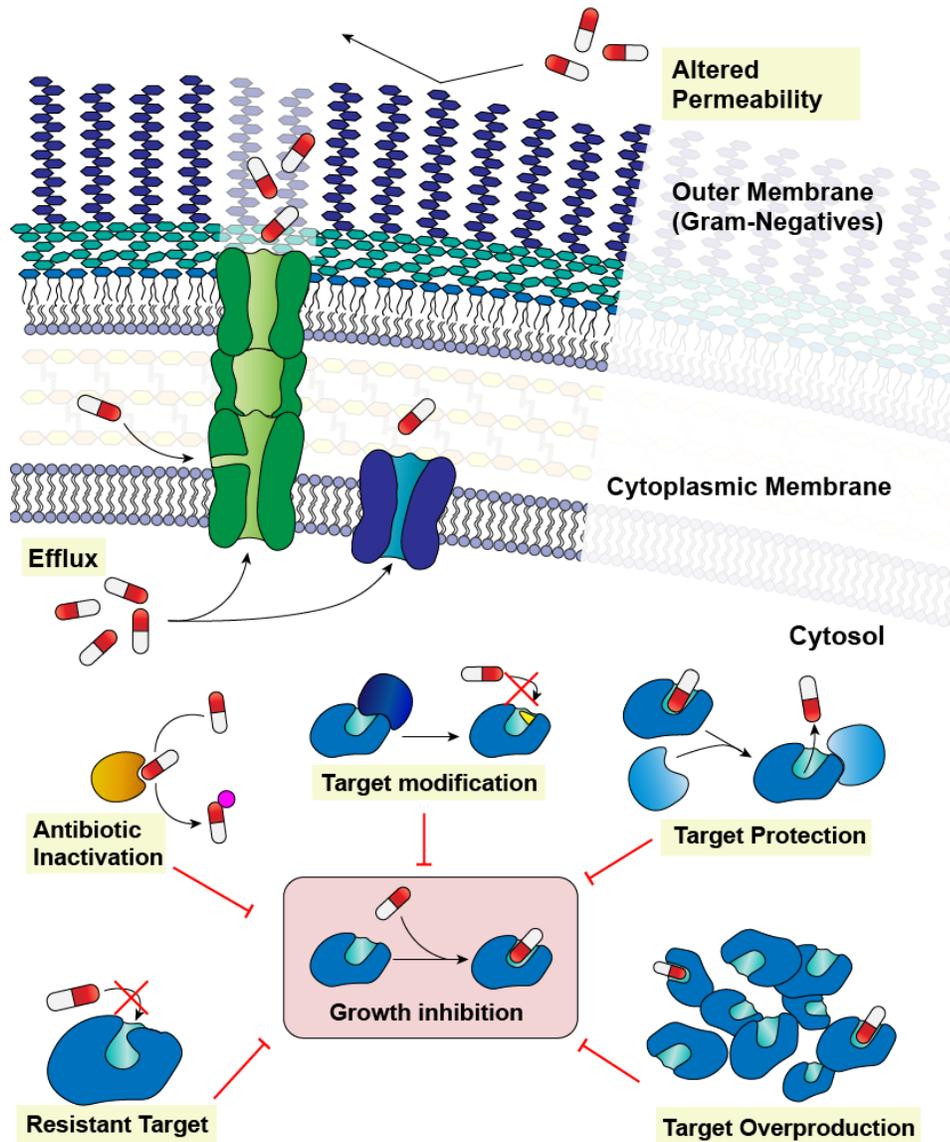
All antibiotic use encourages the development of resistance; therefore, these essential medicines will lose effectiveness the more they're used. Antibiotics are so spectacularly effective and, for the most part, safe that they are used excessively and unnecessarily in people<sup>30</sup>. Over the past few decades, stewardship efforts have become common in many countries, aiming to minimize unnecessary or inappropriate antibiotic use and curtail resistance<sup>31</sup>. However, inappropriate prescription of antibiotics is still rampant. For example, in the US, it is estimated that 30-50% of antibiotic prescriptions in outpatient settings are unnecessary. In Ontario, Canada, estimates suggest that from 2011

to 2015, almost one in five patients received an antibiotic when they saw a physician for the common cold<sup>32,33</sup>. Additionally, antibiotics have found extensive use in agriculture. When added to feed, they can increase the size of livestock and allow for higher population density in farms. Despite banning their use as growth promoters in many countries, the quantity of antibiotics used in agriculture globally is still staggering, at almost 100 000 tonnes<sup>34,35</sup>. Our overuse and misuse of antibiotics have fueled the age of resistance.

In the 20<sup>th</sup> century, rising resistance was offset by the discovery of new antibiotics and the derivatization of existing molecules to evade resistance. For various reasons, some scientific and some economic, this is no longer the case<sup>18,36</sup>. The pipeline for new antimicrobial agents has effectively run dry, further exacerbating the antibiotic resistance crisis. The future will require a multi-pronged approach consisting of new antibiotics and antibiotic combinations coupled with alternative anti-infective strategies such as antibiotic adjuvants, vaccinations, bacteriophages, bacteriolytic enzymes, nanomaterials, etc<sup>37</sup>.

### **Mechanisms of resistance**

A comprehensive understanding of antibiotic resistance at the biochemical level is crucial in combatting resistance because it can help guide the design of molecules that evade specific mechanisms. Antibiotics inhibit bacteria by binding to and interfering with essential cellular processes or structures, such as protein synthesis or the construction and maintenance of peptidoglycan. Bacteria become resistant by preventing, bypassing, or decreasing the strength of these antibiotic-target interactions through a handful of biochemical mechanisms (**Figure 1**)<sup>38</sup>.



**Figure 1 Mechanisms of antibiotic resistance**

For protein targets, mutations leading to amino acid changes in the drug binding pocket can decrease affinity for the antibiotic and confer resistance. Rifamycins, fluoroquinolones, and trimethoprim are all affected by this mechanism. Bacteria can develop these mutations spontaneously or acquire a resistant copy of the target through

horizontal gene transfer<sup>39-41</sup>. Cells can also increase the copies of a target protein present in each cell through mutations that increase their expression or through gene duplication/amplification events. Overproduction of the target requires more antibiotic to inhibit all copies, thereby conferring resistance<sup>42</sup>. For structures that cannot rapidly change their structure through mutation, such as membrane and cell wall components, specific enzymes modify the target to confer resistance. Colistin, for instance, interacts with lipopolysaccharides specific and essential to the Gram-negative cell envelope. MCR-1 transfers a phosphatidylethanolamine group to LPS, which interferes with colistin binding and thus confers resistance<sup>43,44</sup>. Aminoglycoside antibiotics bind to regions of the ribosome composed of RNA, and specific methylating enzymes can modify essential RNA bases to block aminoglycoside binding while conserving ribosome function<sup>45</sup>. Next, target protection proteins confer resistance by physically dislodging antibiotics from their binding pocket. This mechanism is primarily restricted to ribosome-acting antibiotics such as tetracyclines and macrolides, but the Qnr proteins, which confer resistance to quinolones, are a notable exception<sup>46-48</sup>. Resistance can also arise from the direct detoxification of antibiotics by enzymes. These modify the chemical structure of the antibiotic such that it can no longer interact with its target. The hydrolysis of  $\beta$ -lactams by  $\beta$ -lactamases and the phosphorylation of macrolides are particularly relevant examples<sup>49</sup>. Lastly, for antibiotics with an intracellular target, resistance can arise by modulating their ability to reach the cytosol or by their transport out of the cell by efflux pumps<sup>50</sup>. Gram-negative bacteria also possess an outer membrane that acts as a significant permeability barrier to many antibiotics<sup>51</sup>. *Pseudomonas aeruginosa* and *Acinetobacter baumannii* both maintain a high

level of intrinsic resistance due to their highly exclusionary outer membrane and impressive repertoire of efflux pumps<sup>52</sup>. Bacteria muster an array of diverse genetic and biochemical mechanisms to resist the inimical effects of antibiotics.

Many antibiotic resistance mechanisms require specific, dedicated enzymes, which in theory, makes them inaccessible to many microbes. Mutations in the antibiotic target, gene amplification events, and the upregulation of existing proteins, such as efflux pumps, can develop in any bacterium and become fixed following antibiotic selection. In contrast, the *de novo* emergence of dedicated resistance enzymes probably takes thousands of years<sup>53</sup>. The widespread acquisition of resistance in pathogens is partly due to the selection and fixation of various mutations, but the acquisition of new resistance genes from other species by horizontal gene transfer (HGT) is a significant driver of antibiotic resistance<sup>54-56</sup>. This process occurs by three main routes. 1) transformation, the direct uptake of DNA from the extracellular environment; 2) transduction, the introduction of foreign DNA through infection with a bacteriophage; and 3) conjugation, where autonomous or semi-autonomous DNA elements such as plasmids or integrative conjugative elements (ICEs) move between neighbouring cells using sophisticated transfer machinery<sup>57</sup>. While all paths for HGT play a role in resistance, the spread of multi-drug resistance plasmids through conjugation is significant, especially in Enterobacteriaceae<sup>58,59</sup>. The mobilization and dissemination of pre-existing genetic determinants facilitate acquired resistance in common pathogens.

## The resistome

The antibiotic resistome consists of all genes that confer resistance when expressed in the proper context. Bacteria are ubiquitous members of all ecosystems on planet earth and harbour astounding genetic diversity. The golden age of antibiotic discovery leveraged this panacea to develop critical medicines. Antibiotic resistance in pathogenic bacteria was recognized and studied almost immediately, the first penicillinase, produced in *E. coli*, was characterized in 1940<sup>60</sup>. The susceptibility status of non-pathogenic and even environmental microbes received far less attention. It is understood now that the environment is a massive reservoir of diverse antibiotic-resistance genes and mechanisms<sup>38,61</sup>. It is important to catalogue and understand the determinants in this reservoir because they can be mobilized into pathogens by HGT. They can also be present in emerging and opportunistic pathogens native to the environment. There are several examples where non-pathogenic microbes have been identified as the source of resistance in pathogens.

In the 1980s, vancomycin-resistant *Enterococci* (VRE) began to appear in hospitals. These bacteria had acquired *van* operons carried on mobile genetic elements, which remodel cell wall components to confer vancomycin resistance<sup>62</sup>. The most common enterococcal *van* operon (VanA type) shows high identity at the nucleotide level (91-94% for *vanRSHAX*, and 83% for *vanY*) to those in *Paenibacillus thiaminolyticus* PT-2B1, suggestive of recent mobilization from a *Paenibacillus* spp<sup>63</sup>. The CTX-M enzymes are extended-spectrum  $\beta$ -lactamases that can cleave most of, if not all, penicillins, monobactams, and cephalosporins<sup>64</sup>. Phylogenetic analysis suggests that the primarily

environmental Enterobacteriaceae genera, *Klyuvera*, are the source of CTX-M enzymes circulating in the clinic. Remarkably, estimates suggest there have been as many as 8 independent mobilizations from various *Klyuvera* spp, into common pathogens<sup>65</sup>. These cautionary tales highlight the possibility and utility of predicting threats from the resistome.

The reason why the environment is such a hotbed for antibiotic resistance is somewhat trivial. Antibiotics have largely been sourced *from* environmental microbes, meaning that bacteria inhabiting these niches are frequently exposed to antibiotics and are under intense selective pressure to develop resistance<sup>54</sup>. Chemical antagonism of other microbes is common in many genera, particularly the filamentous Actinobacteria mined extensively during the 1930s-1980s. Indeed, large-scale screening of such organisms shows widespread antimicrobial activity; Emerson *et al.* reported antagonism against Gram-positive bacteria and fungi in ~50% of isolates, with lower rates of ~5-20% against Gram-negative species<sup>66</sup>. These estimates align with published systematic reports and our experience with our in-house strain collection (unpublished data)<sup>67</sup>. These figures allude to frequent and intense interbacterial chemical competition in natural environments. Bacteria that can evolve mechanisms to resist this chemical assault have a clear advantage in these ecosystems.

If the production of antibiotics is widespread, then it should follow that resistance is similarly abundant in soil ecosystems. The first large survey of antibiotic resistance in environmental isolates took place in 2006 and sampled 480 actinomycete bacteria for resistance against 21 antibiotics<sup>68</sup>. On average, strains were resistant to 7-8 antibiotics, and some were resistant to as many as 15. Furthermore, resistance in this study was defined as

growth on media containing 20 µg/mL of antibiotic, a relatively high concentration for many antibiotics, and may therefore miss many intermediate or low-level resistance phenotypes. Culture-dependent and culture-independent methods have reiterated the widespread and diverse nature of antibiotic resistance in environmental bacteria<sup>69–75</sup>. Functional metagenomic studies that clone and express libraries of fragmented environmental DNA in *E. coli* directly demonstrate that mobilization into pathogens is possible.

Most accounts of the history of antibiotics are highly anthropocentric. The reality is that antibiotics are ancient molecules. They are exquisitely refined poisons of central processes of the bacterial cell, sculpted by evolution over staggering timescales. Although the precise ecological role of antibiotics is hotly debated<sup>54,76</sup>, their remarkable ability to specifically target essential enzymes, structures, and metabolic pathways within bacteria results from intense natural selection on the interaction between antibiotics and their respective targets. A key but often overlooked contributor to the prevalence of antibiotic resistance in the environment is how long *microbes* have used antibiotics. Several studies have shown that antibiotic resistance pre-dates the anthropogenic use of these compounds, including the recovery and functional validation of resistance genes from 30,000-year-old arctic permafrost and the prevalence of antibiotic resistance in microorganisms from cave systems isolated from the surface world for ~4 million years<sup>75,77,78</sup>. More recently, phylogenetic reconstruction of the biosynthetic pathway of glycopeptide antibiotics has shed light on the ancient history of these metabolites. The analysis of Wagleichner *et al.* places the emergence of glycopeptide antibiotics at 150-400 million years ago; the more

recent end of that estimate is contemporary with the first flowering plants<sup>79</sup>. Remarkably, some antibiotics may be far older. For instance, Hall and Barlow have estimated using phylogenetic reconstruction that Class A  $\beta$ -lactamases emerged ~2.4 billion years ago<sup>80</sup>—implying that microbes first synthesized  $\beta$ -lactams *before* this date. Not only are antibiotics natural molecules produced by ubiquitous environmental genera, but it seems likely that most classes of antibiotics have existed for millennia. The ubiquity and age of antibiotics offer a straightforward explanation for the remarkable diversity and complexity of resistance mechanisms employed by bacteria.

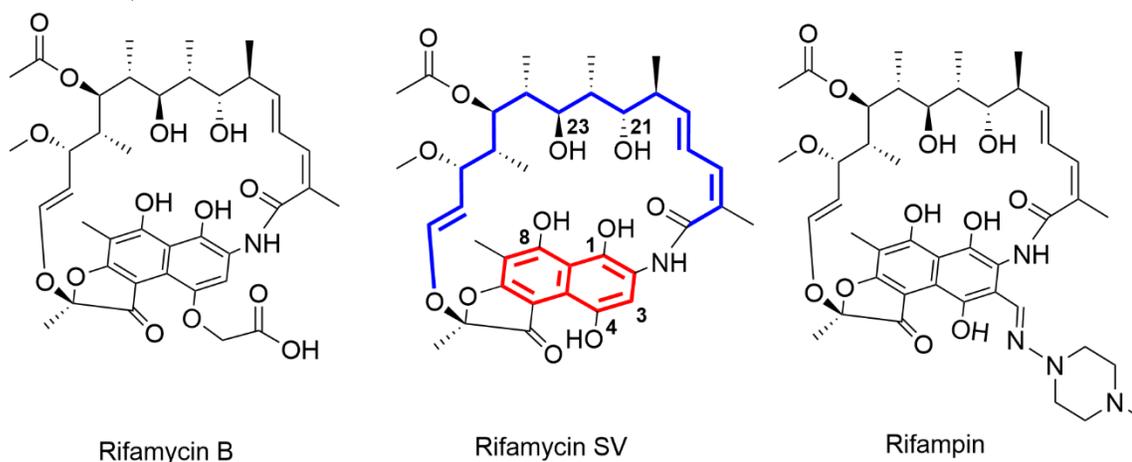
## RIFAMYCIN ANTIBIOTICS

### Discovery and development of rifamycins

Like so many of our antibiotics, Rifamycins are a product of the Waksman platform. The team of Piero Sensi at the Italian pharmaceutical company Lepetit was the first to isolate and study rifamycin antibiotics. They had isolated a bacterium that would eventually come to be known as *Amycolatopsis mediterranei* (previously *Nocardia mediterranei*, and occasionally referred to as *Streptomyces mediterranei*) whose culture filtrate showed promising antibiotic activity<sup>81,82</sup>. The etymology of the rifamycins isn't derived from Greek or Latin but is instead a reference to the French heist movie Rififi, which was popular among staff at Lepetit and became a nickname for this antimicrobial lead. Fermentations of *A. mediterranei* yielded a mixture of related rifamycins (A through E), but rifamycin B, a modestly active antibiotic, was the only compound successfully isolated. Soon after, they realized that rifamycin B transforms into the highly potent rifamycin SV in aqueous solutions. Rifamycin SV is the prototypical natural product

rifamycin and possesses excellent activity against Gram-positive organisms, including mycobacteria (**Figure 2**)<sup>83</sup>. Rifamycins belong to the ansamycin family of compounds that derive their name from the Latin *ansa*, meaning handle, in reference to their aromatic core bridged at non-adjacent carbons by a long aliphatic chain. Rifamycin SV was used to treat some Gram-negative and Gram-positive infections in some countries but had poor pharmacological properties, poor oral bioavailability, and rapid excretion by the liver when administered intravenously<sup>82</sup>. To improve upon these shortcomings, Lepetit generated dozens of rifamycin analogs and revealed the structure-activity relationship of the rifamycins<sup>82,84</sup>. The rules for rifamycin activity, as stated by Sensi, are as follows (**Figure 2**):

- 1) Modifications to the ansa chain decrease activity
- 2) O21 and O23 on the ansa chain are required for activity
- 3) O8 hydroxyl group is required for activity
- 4) Naphthoquinone system not required for activity but O1 is required (either as hydroxy or carbonyl)
- 5) Conformation of the ansa chain cannot be altered
- 6) Substitutions at C3 and C4 are well tolerated



**Figure 2 Structure and SAR of rifamycin antibiotics** Important features of rifamycins are shown on Rifamycin SV. The ansa chain of rifamycins is depicted in blue, the naphthoquinone core is shown in red, and important carbon residues are numbered.

The standout molecule produced by Lepetit's medicinal chemistry program, rifampin (also known as rifampicin), was approved in Italy in 1968 and subsequently in the US in 1971, five years after its discovery (**Figure 2**). Rifampin was orally bioavailable while retaining the potent antimicrobial activity of rifamycin SV<sup>82</sup>. Rifampin soon emerged as a cornerstone treatment for tuberculosis. For this reason, it remains a World Health Organization (WHO) essential medicine to this day<sup>85</sup>.

### **Use of rifampin in tuberculosis therapy**

*Mycobacterium tuberculosis*, the causative agent of tuberculosis, has afflicted human populations for thousands of years. Tuberculosis is primarily a pulmonary disease causing wasting and severe damage to the lungs over time; left untreated, it has a high mortality rate<sup>85</sup>. It can be spread from person to person by droplets and aerosols, aided by the persistent and chronic cough, which is a hallmark of the disease. Tuberculosis was known to Hippocrates, who lived in the 3<sup>rd</sup> and 4<sup>th</sup> centuries BC, and tuberculosis pathology is present in Egyptian remains dating back to 2400 BC<sup>86</sup>. At the beginning of the 20<sup>th</sup> century, it was still a leading cause of death across much of Europe and North America. In the 21<sup>st</sup> century, tuberculosis has been almost eradicated in the developed world but is still a significant problem in areas experiencing significant poverty<sup>87</sup>. In 2021 the WHO estimates that *M. tuberculosis* caused 1.6 million deaths and 6.4 million new infections<sup>85</sup>.

Unique aspects of *M. tuberculosis* physiology and pathology make it a difficult pathogen to treat with antimicrobials. Firstly, mycobacteria possess an unusual cell envelope. The outermost layer consists of long, highly saturated fatty acids (mycolic acids) tethered to the cell by branching arabinogalactan polymers. The mycolic acids form a

second membrane, the mycomembrane, analogous to the outer membrane of Gram-negatives but far more hydrophobic<sup>88</sup>. Secondly, *M. tuberculosis* can enter prolonged periods of metabolic dormancy where cells are not actively replicating but remain viable to resume growth later<sup>89</sup>. Many people harbour a small population of dormant *M. tuberculosis*, referred to as latent tuberculosis. Latent infections are asymptomatic but carry a small risk of progressing to active disease, and when detected, they are usually treated out of an abundance of caution. Not all antibiotics are capable of killing cells in the dormant state.<sup>90</sup> Lastly, *M. tuberculosis* survives and replicates within alveolar macrophages, which typically engulf and destroy bacteria in the lungs. Infected macrophages travel from the airways deeper into the pulmonary tissue and become encased by a network of immune cells attempting to contain the infection, forming what are known as solid granulomas<sup>89</sup>. For antimicrobials to be effective against *M. tuberculosis*, they need to be able to cross the mycomembrane, kill both active and dormant cells, and penetrate macrophages and granulomas. Rifamycins are among the select few which can accomplish all these feats.

The development of successful treatments for tuberculosis has relied heavily on polypharmacy and long treatment courses. Early results with streptomycin, para-aminosalicylic acid, and isoniazid showed promising initial results but, when administered individually, caused frequent relapse due to the development of resistance<sup>91</sup>. A combination of all three agents was far more effective and widely used until rifampin was introduced in the 1960s. This treatment took 1.5 to 2 years, including streptomycin injections for the first 6 months<sup>92</sup>. It is hypothesized that a longer treatment duration is required to kill all dormant *M. tuberculosis*, consistent with clinical data showing that

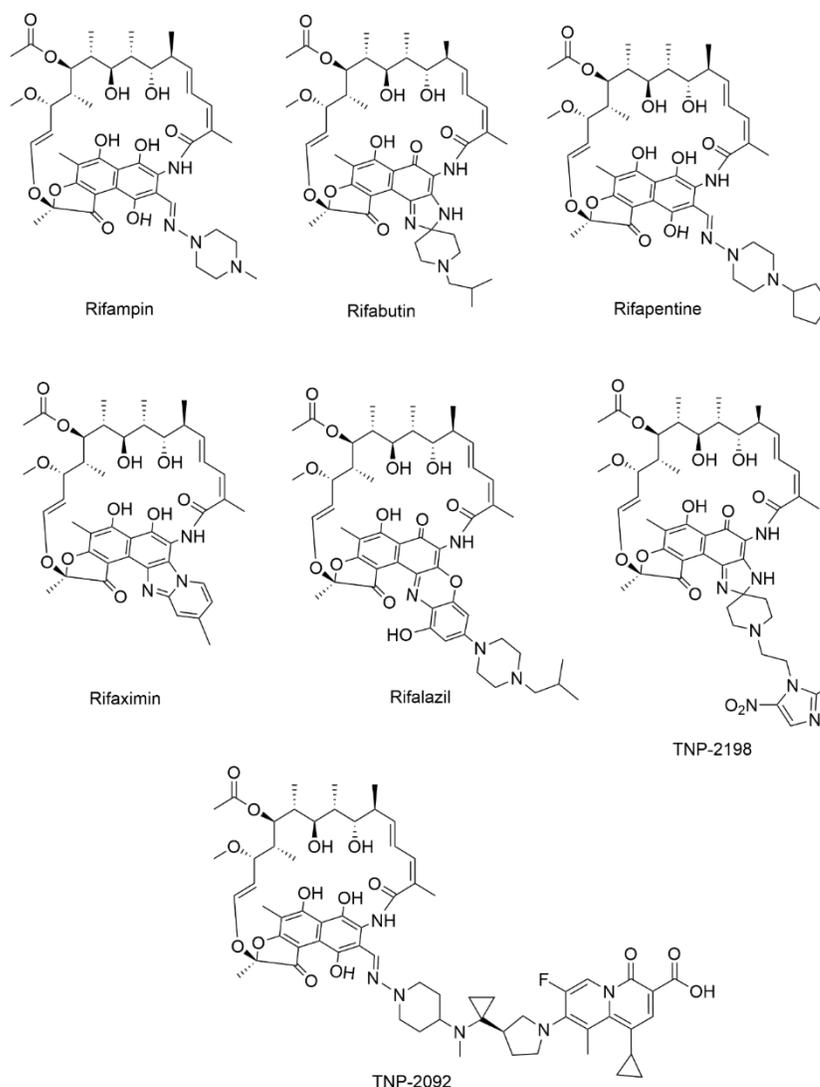
shorter courses often result in relapse. The introduction of rifampin was transformative, offering material improvements in cure rate, halving the length of therapy (to 9 months), and a transition to all oral medicines, namely rifampin, isoniazid, and ethambutol<sup>93</sup>. In 1979, adding pyrazinamide further shortened the treatment to 6-8 months and solidified the modern treatment for drug-susceptible tuberculosis. 2 months of daily isoniazid, pyrazinamide, ethambutol, and rifampin followed by 4 months of rifampin and isoniazid<sup>94</sup>. Rifampin-resistant tuberculosis isolates require entirely different treatment regimens, which have historically been far less effective<sup>85</sup>.

### **Other uses of rifampin**

Rifampin is mainly used to manage tuberculosis but is applicable in several other contexts. For instance, it is the front-line treatment for several non-tuberculosis mycobacteria, such as the slow growers *Mycobacterium kansasii* and *Mycobacterium leprae*, the latter of which is the causative agent of leprosy<sup>95,96</sup>. Rifampin is occasionally used for infections caused by Gram-positive organisms such as *S. aureus* and *Streptococcus pneumoniae*<sup>97</sup>. The addition of rifampin to existing antimicrobial treatments for severe *S. aureus* infections, such as prosthetic joint infections, is common but is becoming more controversial<sup>98-100</sup>. In the case of *S. aureus* bacteremia, a large randomized, double-blind, placebo-controlled trial for adding rifampin on top of standard antimicrobial therapy showed no benefit<sup>101</sup>. Lastly, rifampin is also used as a prophylactic to prevent the spread of bacterial meningitis within the household<sup>102</sup>.

### The next generation rifamycins- rifapentine, rifabutin, and rifaximin

Medicinal chemistry continues to iterate on the rifamycin scaffold, and several next-generation rifamycins have entered the clinic since rifampin in the 1970s. Rifabutin is another semisynthetic rifamycin heavily modified at the C3 & C4 positions; the FDA approved this drug in 1992 for the treatment of *Mycobacterium avium-intracellulare* complex (MAC) in patients with HIV (**Figure 3**)<sup>103</sup>. Rifampin is a potent activator of P450



**Figure 3 Structures of semisynthetic rifamycins**

enzymes in the liver, accelerating the metabolism of many drugs. It is known to lower the effectiveness of anticoagulants, contraceptives, methadone, barbituates, and, critically, several antiretroviral medicines used to manage HIV<sup>104,105</sup>. Patients with HIV are predisposed to various mycobacterial infections such as MAC, *Mycobacterium kansasii*, and even *M. tuberculosis* which all require simultaneous rifamycin and antiretroviral therapy<sup>103</sup>. Rifabutin addresses this critical shortcoming of rifampin; it still induces P450 enzymes but to a far lesser extent than rifampin. As a result, rifabutin is the first-line rifamycin for patients who cannot receive rifampin due to drug-drug interactions.

Rifapentine was approved in 1998, the first new rifamycin approved for *M. tuberculosis* since the introduction of rifampin almost 30 years prior. Rifapentine has a considerably longer half-life than rifampin (14-18 hours vs. 2-5 hours) and was initially approved as a once-per-week treatment of latent tuberculosis (**Figure 3**)<sup>106</sup>. Since its introduction, rifapentine hasn't supplanted rifampin for treating active tuberculosis. This is likely to change following the results of a recent phase 3 clinical trial, which demonstrated the non-inferiority of a four-month course of rifapentine, ethambutol, and moxifloxacin to the previous standard six-month treatment with rifampin, isoniazid, ethambutol and pyrazinamide<sup>107</sup>. Shortening the duration of these regimens, which require monitored daily dosing schedules, is anticipated to limit treatment side effects and minimize unnecessary antibiotic use.

Rifaximin is an orally non-absorbable rifamycin derivative (**Figure 3**). While poor oral bioavailability was a hurdle during the development of rifampin, non-absorbable antibiotics are uniquely well-suited for treating intestinal infections. In 2004 the FDA

approved rifaximin for the treatment of traveler's diarrhea<sup>108</sup>. It has since found widespread use in irritable bowel syndrome<sup>109</sup>.

### **Next generation rifamycins**

Several rifamycin conjugates are currently under development by TenNor therapeutics (China), designated TNP-2092 and TNP-2198 (**Figure 3**). These compounds feature another pharmacophore covalently linked to a rifamycin through the substitution-tolerant C3/C4 region. TNP-2092 is a rifamycin-fluoroquinolone conjugate that retains activity against RNA polymerase (rifamycins), DNA gyrase (fluoroquinolone), and DNA topoisomerase IV (fluoroquinolone). It has a low spontaneous resistance rate and performs better than the co-administration of a rifamycin and fluoroquinolone<sup>110</sup>. This compound has completed phase II clinical trials, and in 2020, the FDA granted it orphan drug status for the treatment of prosthetic joint infections<sup>111</sup>. TNP-2198 contains a nitroimidazole moiety, which is converted into highly reactive radical species by reductases that are used by bacteria, and some parasites, under anaerobic and microaerophilic conditions<sup>112</sup>. These reactive molecules are thought to irreversibly damage DNA and RNA, leading to cell death<sup>113</sup>. TNP-2198 is currently in phase two trials for the treatment of the gastric pathogen *Helicobacter pylori*<sup>111</sup>.

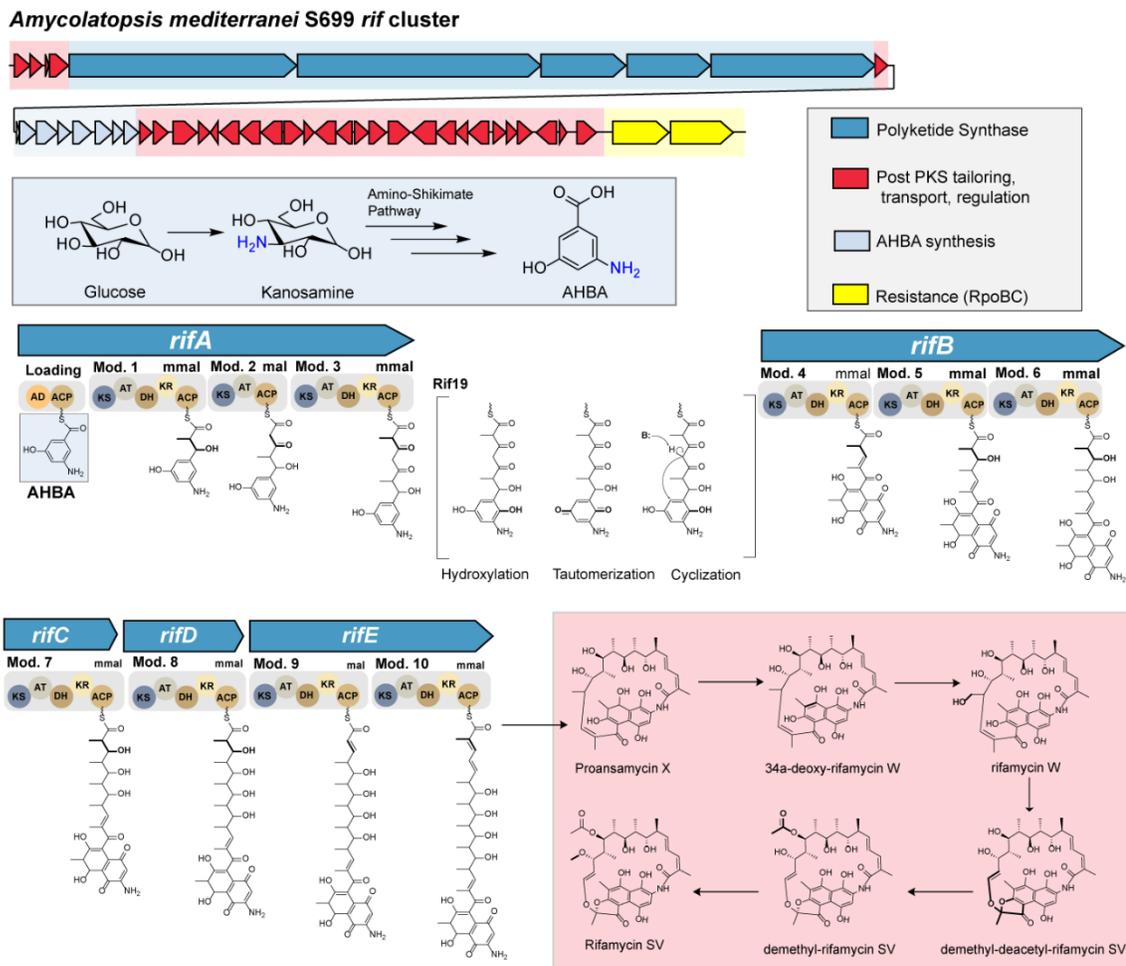
Benzoxazinorifamycins continue the tradition of building complexity outward at the C3/C4 position, resulting in even higher binding affinity and potency<sup>114</sup>. Rifalazil was a benzoxazinorifamycin that was under development for treating *M. tuberculosis*, then *Clostridium difficile* infections, but has been abandoned due to severe side effects several times<sup>115,116</sup>. The benzoxazinorifamycin core of rifalazil is still being studied, with

derivatives reported in 2022 that have improved activity against rifampin-resistant *M. tuberculosis* and decreased induction of P450 enzymes<sup>116,117</sup>.

Unsurprisingly, significant efforts have been made to design rifamycins to overcome common resistance mechanisms. These will be covered later, alongside their corresponding mechanisms of resistance.

### **Rifamycin Biosynthesis**

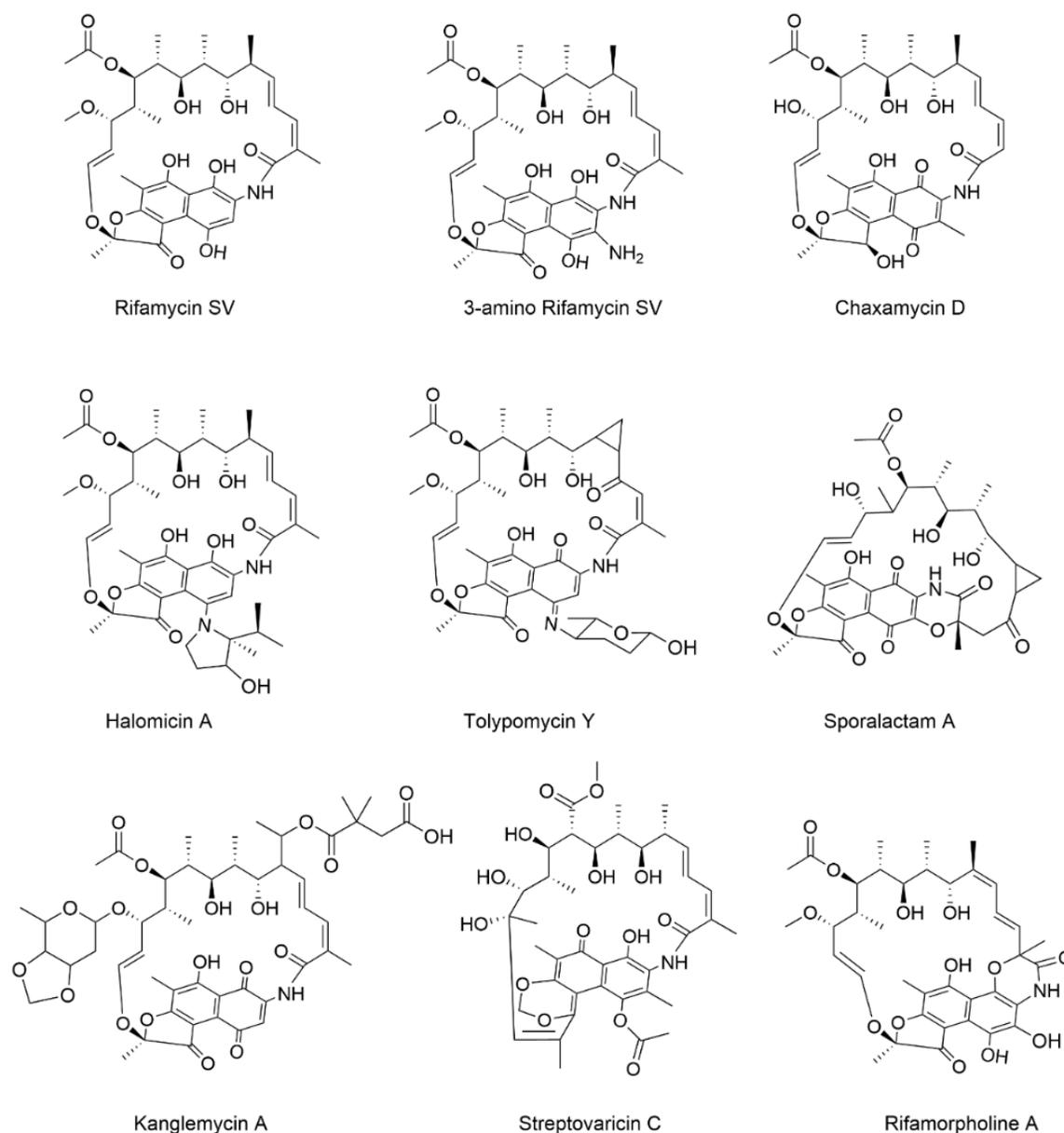
Rifamycins are remarkably complex molecules characterized by a large macrocycle and many chiral centers on their ansa chain; their biosynthesis by microbes is accordingly complex. The biosynthetic genes for rifamycins are arranged in a single ~100kbp locus in *A. mediterranei*, which at its border contains a rifamycin-resistant copy of the  $\beta$ -subunit of RNA polymerase for self-resistance<sup>118</sup>. Conceptually, the biosynthesis can be simplified into three general steps, synthesis of the 3-amino-5-hydroxy-benzoic acid (AHBA) precursor, extension by the modular polyketide synthase (PKS), cyclization, and post-PKS modifications (**Figure 4**)<sup>119</sup>. AHBA is synthesized by a suite of enzymes carrying out a parallel shikimate pathway on an amino-modified precursor (sometimes called the amino-shikimate pathway)<sup>120</sup>. AHBA then serves as the starter unit for the *rif* PKS machinery. This molecule is passed processively down the assembly line, and each of the 10 modules in the *rif* PKS extends the molecule in increments of two carbons using malonyl or methylmalonyl-CoA extender units. Next, the action of specific domains (ketoreductase and dehydratase) instills specific functional groups, double bonds, etc. The biochemistry of PKSs is highly complex and beyond the scope of this chapter; interested readers should consult the following reviews by Floss and Yu or Fischbach and Walsh<sup>119,121</sup>. Rifamycins



**Figure 4 Proposed biosynthesis of rifamycin antibiotics. (Top)** Arrangement of the biosynthetic gene cluster and the synthesis of the 3-amino-5-hydroxybenzoic acid (AHBA) starter unit. **(Bottom)** Iterative synthesis of rifamycins by polyketide synthase machinery (PKS) and the post-PKS processing steps. Individual domains are labelled in the PKS modules. Bioactive rifamycins are demarcated by the transition from rifamycin W, and dimethyl-deacetyl-rifamycin SV, corresponding to the synthesis of the 5-membered ring. KS, ketosynthase, ACP, acyl carrier protein, DH, dehydratase, KR, ketoreductase, AD, adenylation. Mal and mmal denote malonyl-CoA and methylmalonyl-CoA extender units, respectively.

contain a naphthoquinone aromatic core, whereas the AHBA starter only contributes a benzyl. The second ring is formed on the PKS, following the third module by an intramolecular reaction catalyzed by Rif19<sup>122</sup>. After all the PKS modules are complete, rifamycins are cyclized through the formation of an amide bond and subject to a host of post-PKS tailoring enzymes required to form rifamycins with any antibiotic activity<sup>119</sup>.

In nature, rifamycins are not a static chemical entity. Rifamycin-like molecules are produced in *Amycolatopsis*, *Nonomurea*, *Micromonospora*, *Salinispora*, *Actinomadura*, and *Streptomyces* spp., spanning 4 different families within the Actinobacteria. Differences primarily in the post-PKS tailoring steps among various rifamycin producers yield diverse compounds such as chaxamycin, tolypomycin Y, halomicin, streptovaricin, aminorifmycins, sporalactams, morpholinorifamycins, and kanglemycins (**Figure 5**)<sup>123–129</sup>. Conceptually it is important to recognize that the rifamycin resistome has been shaped by all these compounds and all their producers, not just rifamycin SV and *Amycolatopsis mediterranei*.



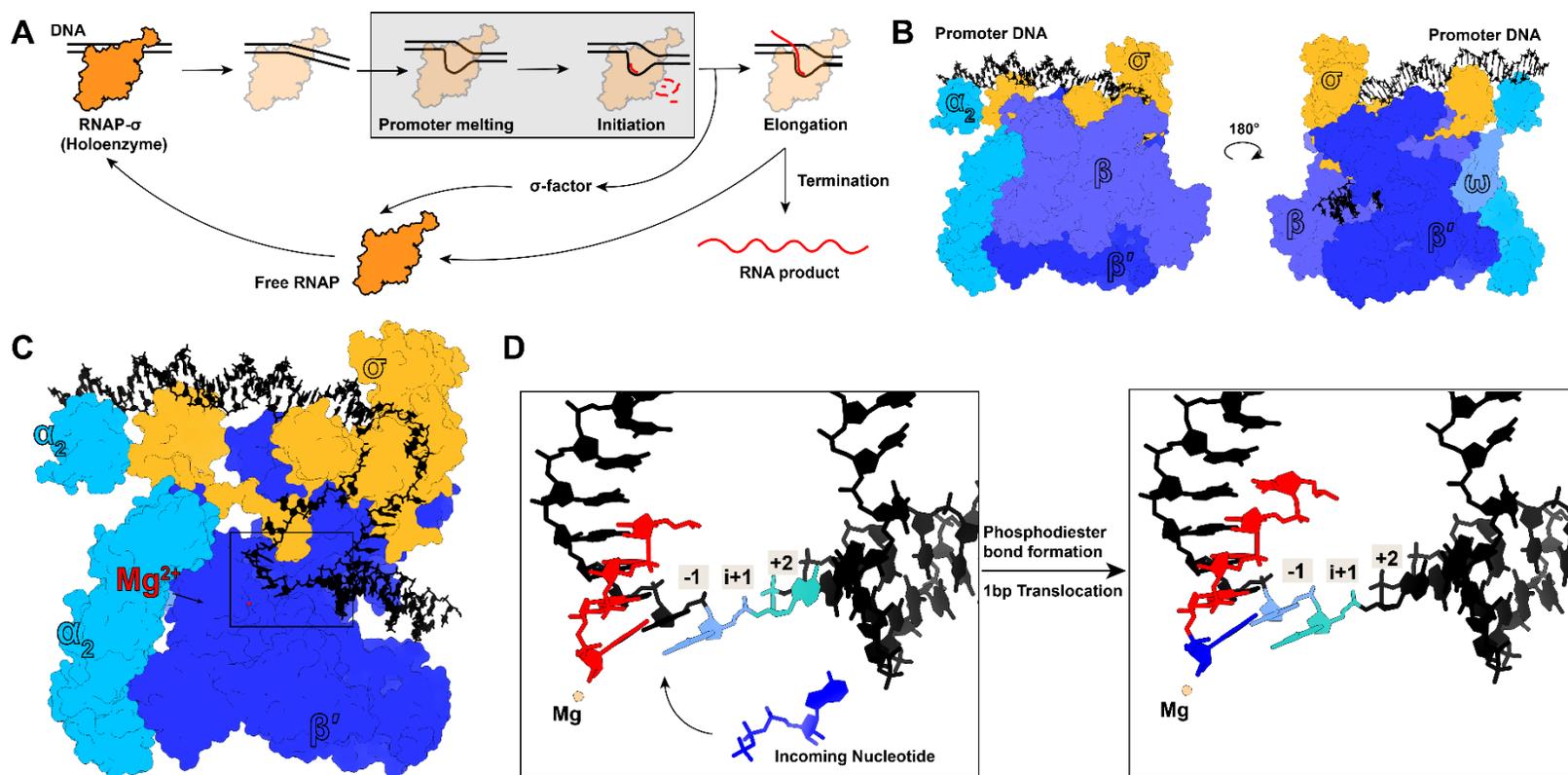
**Figure 5 Diversity in natural product rifamycins**

### Prokaryotic RNA polymerase

Rifamycins attain their potent antimicrobial activity by selectively inhibiting prokaryotic DNA-dependent RNA polymerase (RNAP). All cellular lifeforms encode at least one RNAP to generate all the RNA species required for life. This includes messenger

RNA, the templates for protein synthesis, and much of the protein synthesis machinery itself. The ribosome consists predominantly of RNA and uses activated amino acids attached to RNA scaffolds to facilitate protein synthesis (transfer-RNAs). Eukaryotic cells delegate the synthesis of these various RNAs to three distinct RNAP isoforms (I-III). Prokaryotes, on the other hand, use a single enzyme for all RNA synthesis, making this enzyme an excellent target for antibiotics<sup>130</sup>

Prokaryotic RNAP is composed of 5 proteins with the following stoichiometry, two  $\alpha$ -subunits, a  $\beta$  subunit, a  $\beta'$  subunit, and the  $\omega$  subunit. These 5 proteins make up the core enzyme which is catalytically competent but requires an additional factor,  $\sigma$ , to initiate transcription (**Figures 6A and B**). RNAP- $\sigma$  complexes (holoenzymes) are directed to precise genomic locations by sequence-specific interactions between DNA and  $\sigma$ -factors<sup>131</sup>. These loci are known as promoters. Transcription begins with the formation of a closed complex, where RNAP is bound to fully dsDNA. Next, RNAP ‘melts’ promoter DNA to separate the template and non-template strands and form the transcription bubble, creating what’s known as the open complex (**Figure 6C**). Now, template DNA is in the active site and primed for Watson-Crick-Franklin hydrogen bonding to incoming ribonucleotide triphosphates (rNTPs). When incoming nucleotides are added to the transcript (i+1 site), RNAP moves forward to bring the next template nucleotide (+2 site) into the active site (i +1 site)(**Figure 6D**)<sup>132</sup>. The isomerization to fully processive, elongating RNAP species requires the loss of the  $\sigma$ -factor, which blocks the RNA exit tunnel and is bound tightly to the promoter. Recent work suggests that the growing transcript displaces the  $\sigma$ -factor. Still, this process isn’t efficient and usually entails the



**Figure 6 The transcription cycle** **A)** Steps of the transcription cycle. **B)** Structure of RNA polymerase holoenzyme in complex with promoter DNA, subunits are colored and labelled accordingly. **C)** RNAP open promoter,  $\beta$ -subunit removed so that the transcription bubble can be viewed. Black rectangle denotes the active site visualized in **D)** Translocation of RNAP following nucleotide addition. Promoter DNA is colored black and RNA is red. The template for the incoming nucleotide (i+1) and the subsequent base (+2) are colored to illustrate the movement of DNA through the active site. All figures were generated using ChimeraX v1.5 using models deposited in the PDB with the following accession codes, 6PSW (**B and C**), 6KOQ(**D**).

iterative synthesis of 8-15bp transcripts until the  $\sigma$ -factor is successfully displaced and the promoter is escaped<sup>133</sup>. This process is analogous to the ‘turning over’ of an internal combustion engine. Elongating RNAP is highly processive, synthesizing RNA at a rate of ~20-40bp per second<sup>134,135</sup>. Lastly, the end of transcription is signalled by specific DNA and RNA sequences which, sometimes in cooperation with termination factors (such as Rho), cause the dissociation of RNAP, DNA, and RNA, thus completing the transcription cycle<sup>136,137</sup>. The mechanistic complexity in transcription has invited the production of antibiotics that target every step of this cycle, most notably but not limited to the rifamycins<sup>138</sup>.

### **Mechanism of rifamycin antibiotics**

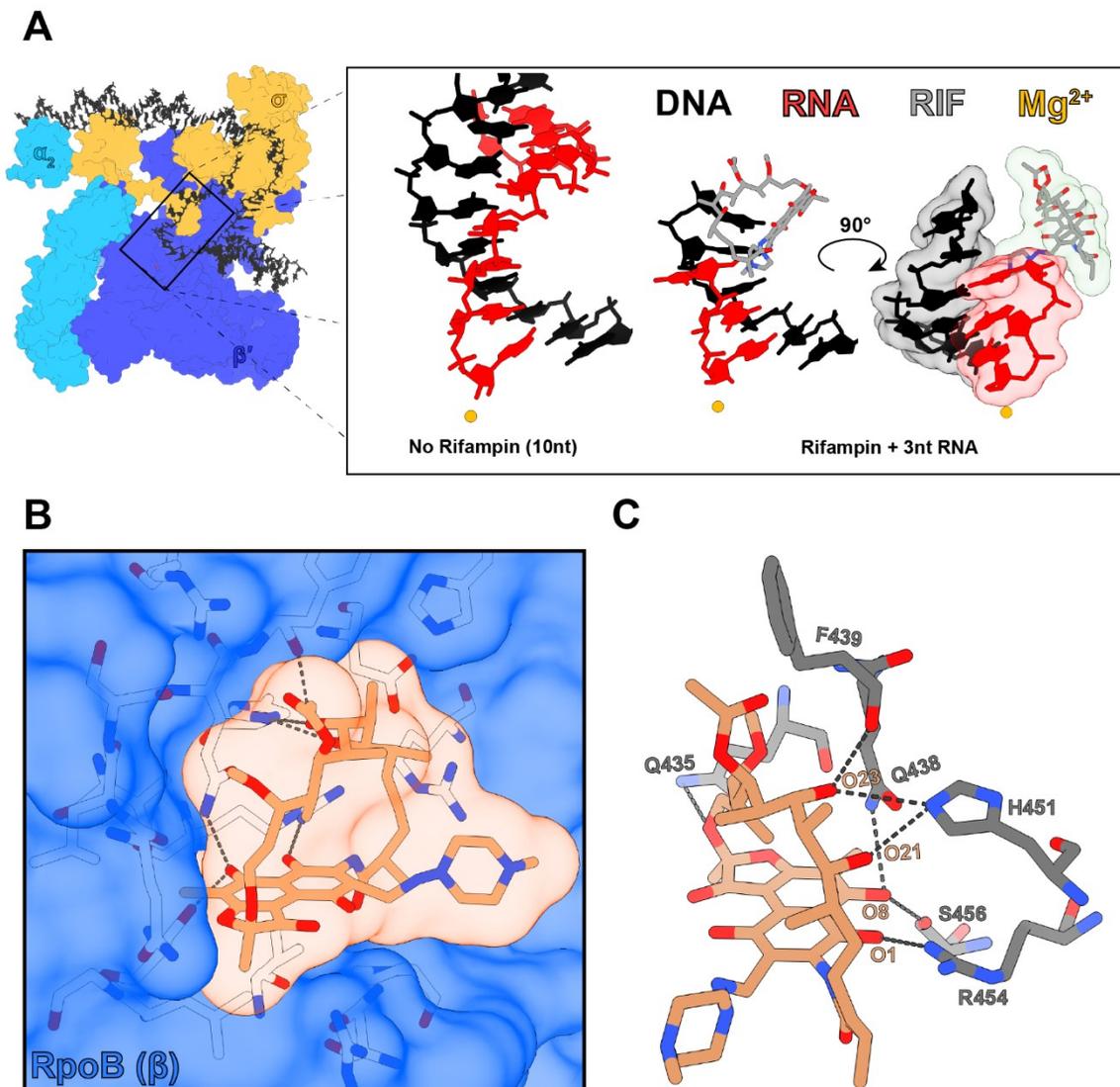
Despite increasingly widespread use in the mid-1960s, how rifamycins inhibited the growth of bacteria was still a mystery. In 1965, researchers in Rome noted that rifamycin almost completely blocked the incorporation of <sup>14</sup>C labelled uridine and leucine into RNA and protein using cell-free extracts of *B. subtilis*<sup>139</sup>. At the time, Calvori *et al.* considered it most likely for rifamycins to act on protein synthesis, and the drop in uridine incorporation was a side effect. The next few years saw the development of methods for purifying and studying bacterial RNAP *in-vitro*, setting the stage for the study of rifamycins to truly begin. In 1967 Hartmann *et al.* showed that rifamycins specifically inhibit bacterial RNAP, and in 1968 Sippel and Hartmann noted that rifamycins only affect RNAP before transcription has begun<sup>140,141</sup>. They correctly identified transcription initiation as the specific process inhibited by rifamycins. Contemporary work also showed spontaneous rifamycin-resistant mutants possessed RNAP that no longer interacted with rifampin<sup>142</sup>.

The modern understanding of the mechanism of action of rifamycins was put forth in the late 1970s, based on seminal experiments by William McClure and Carol Cech<sup>143</sup>. They performed *in-vitro* transcription reactions where they could limit the transcript size to only 2 or 3nt long by withholding required nucleotides. This enabled them to directly measure the effect of rifampin on phosphodiester bond formation. They found that rifampin had only a very modest impact on the formation of the first phosphodiester bond (dinucleotide production). In stark contrast, rifampin completely inhibited the synthesis of trinucleotide product. In the presence of all four NTPs, dinucleotide synthesis continued unabated, but RNAP could not produce longer species. This futile production of dinucleotides is called abortive transcription. Outside of very few exceptions, an NTP (i.e., pppN) initiates transcription in bacteria. Therefore, the dinucleotide formed in the first phosphodiester bond possesses a 5'triphosphate cap (pppNpN). Because the 5'triphosphate cap isn't required for phosphodiester bond formation, RNAP accepts various initiating nucleotides *in vitro*. Remarkably, when NDP, NMP, or a 5'OH dinucleotide (NpN) was used to initiate transcription, it released the inhibition of the second phosphodiester bond, and 3nt transcripts, ppNpNpN, pNpNpN, or NpNpN could now be formed in the presence of rifampin. To McClure and Cech, this could be rationalized by a steric block imposed by rifamycins during the translocation step, which normally follows the formation of the first phosphodiester bond. They hypothesized this because inhibition of transcription seemed to occur once the nascent RNA had reached a particular size. They summarized their theory, the steric occlusion model, in an eerily prescient figure which I've reproduced in **Figure 7**.

DNA Template	---C	G	T	A	C	A	T---	
	RIF	{			U	G	U	NTP incorporated
Initiating nucleotides	pppA		+	-	-			
	ppA		+	+	-			
	pA		+	+	?			
	CpA		+	-	-			

**Figure 7. The steric occlusion model.** Length of the transcript made by rifampin inhibited RNAP utilizing different initiating nucleotides. CpA (5'OH cytidine-adenine dinucleotide) Adapted from McClure and Cech<sup>143</sup>.

Almost 40 years later, the crystal structure of rifampin in complex with RNAP from *Thermus aquaticus* would cement the model put forward by McClure and Cech<sup>144</sup>. Rifamycins bind to the  $\beta$ -subunit of RNAP (RpoB), ~8 angstroms away from the active site of RNAP, the site of phosphodiester bond formation. Rifamycins bound to RpoB extend into the space occupied by the ~4<sup>th</sup> RNA nucleotide in the nascent RNA:DNA hybrid, forming, as predicted by McClure and Cech, a steric barrier to RNA synthesis. Conversely, a mature transcript would occupy this space and instead occlude rifamycins, the reason that rifamycins cannot inhibit transcription that has already begun. In 2017 Lin *et al.* sequentially crystallized RNAP with promoter DNA, rifampin, and 1, 2, 3, and 4nt RNA (5'OH) products, capturing atomic evidence of the steric occlusion model (**Figure 8A**)<sup>145</sup>. 1-3nt RNAs can co-crystalize with rifampin, but 4nt RNA cannot. The 5'nt closest to rifampin in the 3nt co-structure is not paired with template DNA and is rotated out of alignment due to steric interference from rifampin.



**Figure 8 Rifamycin inhibition of RNA polymerase.** **A)** Steric occlusion of the growing transcript by rifampin. Inset shows DNA-RNA hybrids formed in the active site with and without rifampin. **B)** The rifamycin binding pocket. The surface of the  $\beta$ -subunit is shown in blue, with rifamycin interacting residues shown in white. Rifampin is shown as surface and sticks. **C)** Summary of important hydrogen bonds formed between rifampin and *M. tuberculosis* RpoB. All figures were generated using ChimeraX v1.4. PDB codes used to generate this figure are as follows. Figure A, 6PSW (Open complex depiction), 6KOQ (10nt nascent transcript), 5UHC (open complex with rifampin and 3nt RNA). Figure B and C, 5UHB.

The rifamycin binding pocket is highly conserved at the amino acid and structural level among prokaryotes; divergence in this region in eukaryotic RNAPs renders them immune. Rifamycins bind with a remarkable affinity to bacterial RNAP with dissociation constants ( $K_d$ ) measured at sub-nanomolar concentrations<sup>146</sup>. A wealth of structural information on this interaction is available as rifampin has been crystalized in *T. aquaticus*, *E. coli*, *Mycobacterium smegmatis*, and *M. tuberculosis* RNAP<sup>135,145,147,148</sup>. As these structures are broadly but not entirely in agreement regarding specific rifampin-RNAP interactions, I will summarize the rifampin-RNAP interactions characterized in the *M. tuberculosis* enzyme<sup>145</sup>.

Rifamycins fit snugly into their binding pocket on RpoB, supported by van der Waals interactions with 15 residues and a network of 7 hydrogen bonds (**Figure 8B and C**). Specific interactions between RpoB and essential rifamycin moieties identified by Sensi's team in the 1960s constitute most of the hydrogen bonding network (**Figure 8C**)<sup>82</sup>. On the ansa chain, the O23 hydroxyl group forms hydrogen bonds with F439 and H451, and the nearby O21 also forms a hydrogen bond with H451. The O1 hydroxyl residue forms a hydrogen bond with R454, and the O8 hydroxyl bonds with S456 and Q438. The remaining hydrogen bond occurs between Q435 and the ester oxygen attached to the 5-membered ring; however, this bond isn't consistent across all structures<sup>147-149</sup>. Furthermore, the C3 and C4 positions are amenable to modification and point away from the binding pocket and into the main channel of RNAP, hence their minimal negative impact on RNAP binding (**Figure 8B**). The precise 3-dimensional structure of rifamycins, including the specific conformation of the ansa chain, is necessary to position these four hydroxyl

residues and generate this network of hydrogen bonds. Rifamycins are an excellent example of the remarkable chemical space microbes exploit for interbacterial competition.

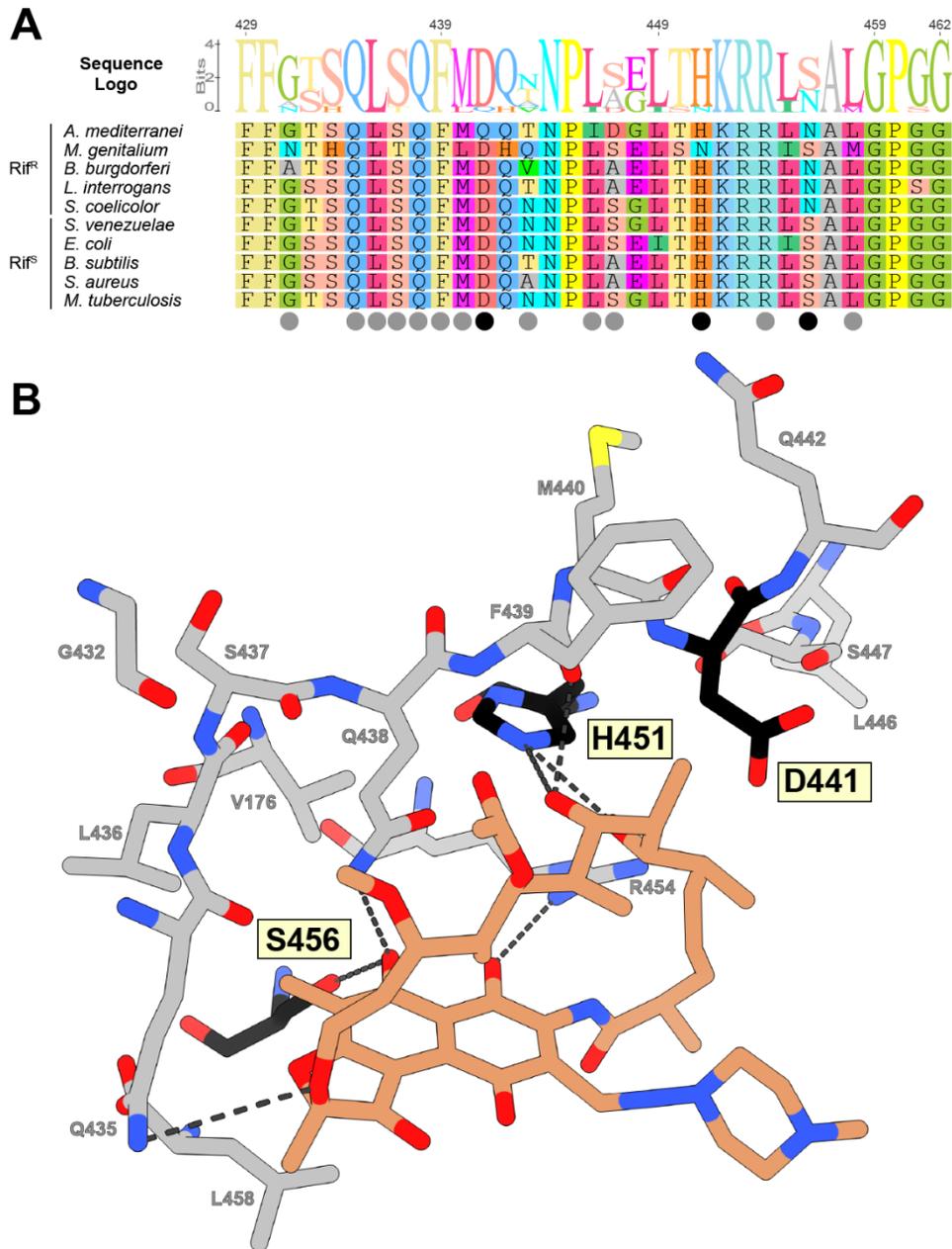
## RIFAMYCIN RESISTANCE

### The Achilles heel – *rpoB* mutations

Rifamycins are highly potent and broadly active, with semisynthetic derivatives possessing good oral bioavailability. Therefore, these compounds should be mainstays in the clinic. Unfortunately, rifamycins suffer from high rates of spontaneous resistance, which has severely limited their use. *In-vitro* selection for rifampin resistance yields a resistant mutant in 1 per  $10^{-7}$  to  $10^{-10}$  divisions, comparable to or smaller than the population size of many infections<sup>150–153</sup>. Rifamycins are relatively unique in that a single locus, *rpoB*, encodes its target. Many classes of antibiotic target multiple related enzymes, for instance, the  $\beta$ -lactams usually inhibit multiple PBPs, and the fluoroquinolones inhibit both DNA gyrase and topoisomerase IV<sup>154,155</sup>. In contrast, translation inhibitors usually interact with RNA portions of the ribosome, and cells encode multiple copies of these RNAs<sup>156</sup>. These factors limit the impact of a point mutation in a single enzyme or RNA. On the other hand, a single mutation in *rpoB* will alter all RNAP within the cell. In addition to being relatively frequent, substitutions in rifamycin binding pocket generally confer a high degree of resistance ( $> 100$ -fold increase in MIC)<sup>157</sup>. Substitutions in RpoB usually decrease fitness due to pleiotropic effects on transcription and gene expression. However, this cost seems insufficient to prevent their survival *in vivo* and subsequent compensatory mutation(s) in core transcription machinery, mitigating their deleterious effects<sup>158–162</sup>. Rifamycin

monotherapy is avoided outside of highly specific exceptions to prevent the selection of resistant alleles by spontaneous mutation.

The ease with which spontaneous rifampin resistance develops was a boon to bacterial geneticists, who could map resistance-conferring mutations to *rpoB* as early as 1970<sup>163</sup>. More detailed maps emerging in the 80s and 90s revealed that >95% of all resistant mutants mapped to a single 81bp region within *rpoB*, termed the rifampin resistance determining region (RRDR), or cluster I in the map developed by Jin and Gross<sup>39,164</sup>. This stretch of amino acids is highly conserved in bacteria (**Figure 9A**). Structural data show that the RRDR encompasses most of the rifamycin binding pocket, so it is logical that substituting residues lining this pocket lower the affinity of RpoB for rifampin. Within the RRDR, hotspots for resistance exist, D441, H451, and S456 (*M. tuberculosis* numbering) account for 9%, 36%, and 41%, respectively, of rifampin-resistant *M. tuberculosis* isolates from a collection of almost 500 (**Figure 9A**). The specific substitution of S456L alone accounts for almost 40%, and H451Y another 20%<sup>165</sup>. These are significant residues because they both form hydrogen bonds with rifamycins; in both cases, critical hydrogen bonds are lost and replaced with bulkier side chains, making RNAP effectively immune (**Figure 9B**). The precise role of D441 is less certain; in the *T. aquaticus* rifampin-bound RNAP, this amino acid appeared to form a hydrogen bond with rifampin through O21; however, in other structures with more susceptible polymerases, this has not been observed<sup>145,148</sup>. Furthermore, a D441N substitution, which would retain the ability to form a hydrogen bond with O21, confers rifampin resistance<sup>166</sup>. More likely, the negatively charged D441 assists rifamycin binding by neutralizing the positive charge from two



**Figure 9** The rifamycin resistance determining region (RRDR) of RpoB. **A)** The RRDR of rifamycin susceptible and resistant bacteria. An alignment of the RRDR with the consensus sequence depicted as a sequence logo above the alignment. Circles below each amino acid represent residues where substitutions confer resistance; black circles represent the three most frequently substituted residues, grey represent infrequent sites. **B)** Structural context of residues associated with rifampin resistance. Residues are color-coded according to frequency, as described above. This figure was generated using ChimeraX v1.4 and PDB model 5UHB.

nearby arginines and contributing van der Waals interactions<sup>144,148</sup>. Although D441, H451, and S456 are the most common locations of resistance substitutions, many others have been characterized, including some outside the RRDR, although these are relatively rare. **Figure 9** depicts sites within the RRDR associated with rifamycin resistance. Apart from acquired resistance during rifamycin treatment, some bacteria host naturally resistant RpoBs. These include medically relevant genera such as *Bordetella*, *Leptospira*, and *Mycoplasma*. *Bordetella* and *Leptospira spp.* have S456N substitutions, whereas *Mycoplasma* possesses H451N (**Figure 9A**). The S456N substitution is also found sporadically throughout the actinomycetes, such as *Streptomyces coelicolor* (**Figure 9A**)<sup>167</sup>. In addition, many *Nocardia spp.* possess a second, rifamycin-resistant (H451N) copy of *rpoB*<sup>168</sup>. Lastly, the producers of rifamycin antibiotics also have resistant RNAPs (D441G, S456N) as a requisite for producing these molecules (*A. mediterranei*, **Figure 9A**)<sup>119</sup>.

The relative homogeneity of rifamycin-resistant *rpoB* alleles led many researchers to search for compounds that bind to resistant polymerases. For instance, some benzoxyaminorifamycin analogs had markedly improved activity against common resistant alleles relative to rifampin. Extra interactions made by the bulky substituents extended from the C3/4 positions with a nearby portion of the  $\sigma$ -factor are thought to be responsible. However, rifalazil itself did not, and the magnitude of inhibition was still sufficiently low so as not to be clinically valuable against resistant isolates<sup>114,116</sup>. Natural products have also yielded several compounds able to bind rifampin-resistant RpoBs. For instance, sorangicin is a structurally unrelated polyketide that remarkably also targets the rifamycin binding pocket. Cells develop resistance to sorangicin at a similar rate, but

intriguingly they show only partial cross-resistance with rifamycins. Significantly, sorangicin retains significant activity against S456L alleles due to differences in how it binds to RNAP<sup>169</sup>. Recent work has suggested subtle mechanistic differences in how sorangicin inhibits transcription against S456L polymerases, and it remains a promising lead compound<sup>170</sup>. Kanglemycins are highly decorated rifamycins first discovered in 1988<sup>123</sup>. 20 years later, two independent groups realized that Kanglemycins retained the ability to inhibit both S456L and D441V alleles with efficacy on par with the wildtype enzyme. By virtue of their decorations, kanglemycins make additional contacts with RNAP which help them bind to RNAP with substitutions in the RRDR. Interestingly, they also extend closer to the active site and, unlike rifamycins, can inhibit the formation of the first phosphodiester bond during transcription<sup>147,171</sup>. Unfortunately, H451Y retains high-level resistance to both sorangicin and kanglemycins. Nonetheless, these compounds may have therapeutic potential, and orally bioavailable kanglemycin derivatives have already been synthesized<sup>172</sup>. The clinical utility of molecules that could inhibit resistant *rpoB* alleles is enormous. Combinations of such compounds could treat existing resistant strains of *M. tuberculosis*, suppress the emergence of resistance, and potentially open the door for broader use of rifamycins.

### **Permeability and efflux**

Rifamycins are often considered Gram-positive specific antibiotics. In truth, they have modest activity against many Gram-negatives, including Enterobacteriaceae. However, the outer membrane poses a significant barrier to these compounds, as mutants with altered outer membrane permeability can be ~100-fold more susceptible<sup>173,174</sup>.

Consequently, rifamycins have excellent synergy with compounds that disrupt the outer membrane, such as polymyxins<sup>175,176</sup>. On the other hand, the efflux of rifamycins by Gram-negatives seems to be a minor contributor to their insensitivity to rifamycins<sup>177</sup>. These intrinsic barriers have prevented the use of rifamycins against most Gram-negative bacteria.

A handful of efflux pumps are associated with low-level rifampin resistance in *M. tuberculosis*. In general, even the overexpression of these pumps confers only low levels of resistance (2-4 fold)<sup>178–180</sup>. Although the cell envelope of *M. tuberculosis* can prevent the accumulation of many antibiotics in its cytosol, rifamycins seem well-suited to bypass these defences.

### **RNAP programmes and RNAP binding proteins**

While all bacteria use a dedicated housekeeping  $\sigma$ -factor to direct RNAP to essential genes such as those encoding RNAP, other  $\sigma$ -factors control more specific transcriptional sub-programmes. Knockouts of  $\sigma^B$ , an alternate  $\sigma$ -factor in mycobacteria, show a modest increase in rifampin susceptibility. Recent work suggests this is because  $\sigma^B$  also directs the production of more RNAP, and that increased susceptibility results from a smaller, and therefore easier to inhibit, pool of RNAP<sup>181</sup>.

A small RNAP accessory protein unique to Actinobacteria, called RbpA, is required for wild-type levels of rifampin resistance. First discovered in *S. coelicolor*, the deletion of *rbpA* sensitized cells to rifampin, leading to several studies investigating its role in resistance and even suggestions that it interfered with rifampin binding<sup>182,183</sup>. Further studies would find that RbpA has a general role in transcription, stabilizing open promoter

complexes to stimulate transcription. More recent work has also refuted that RbpA affects rifampin's ability to inhibit RNAP<sup>184-186</sup>. Like  $\sigma^B$ , the stimulation of transcription provided by RbpA may be an essential determinant of RNAP levels in the cell.

Neither of these mechanisms is known to cause rifamycin treatment failure and are perhaps best conceptualized as intrinsic resistance mechanisms. However, this could be debated on the grounds that both are present in rifampin-susceptible *M. tuberculosis*.

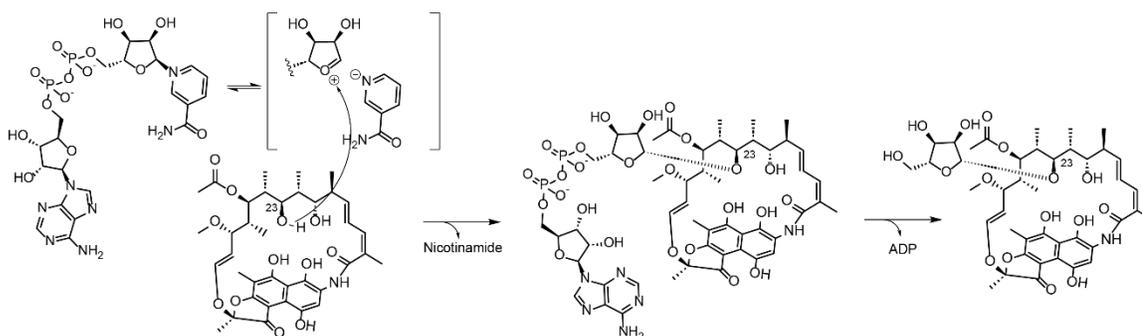
### **The environmental rifamycin resistome**

Investigation of rifamycin resistance outside the clinical paradigm of *rpoB* alleles has unearthed a surprising variety of enzymatic inactivation mechanisms. Dedicated resistance enzymes which catalyze the ADP-ribosylation, glycosylation, phosphorylation, and monooxygenation of rifamycin antibiotics have all been described<sup>187</sup>. Rifamycin inactivation is found in *Streptomyces*, *Nocardia*, *Rhodococcus*, *Mycobacterium*, and some Firmicutes such as *Bacillus* and *Paenibacillus*. The mechanistic diversity of these enzymes rivals or surpasses those that exist to counter any other class of antibiotic. The evolution of many routes to detoxification suggests that rifamycins are an ever-present threat and that perhaps developing and maintaining rifamycin-resistant RNAP is untenable for many bacteria. The investigation of the rifamycin resistome demonstrates why this type of research is necessary; it has revealed many widely disseminated genes that may threaten these essential medicines' efficacy.

### **ARR**

Dabbs *et al.* first described the accumulation of inactive ribosylated rifamycins in culture broths of fast-growing mycobacteria in the mid-1990s<sup>188,189</sup>. Subsequently, the gene

responsible, termed *arr*, was discovered in *M. smegmatis* (*Mycobacterium smegmatis*), and the nature of the enzymatic mechanism was further clarified<sup>190,191</sup>. Arr uses NAD<sup>+</sup> as a substrate to transfer ADP-ribose to the O23 hydroxyl group of rifamycins, releasing nicotinamide in the process (**Figure 10**)<sup>192</sup>. ADP-ribosylated rifamycin species subsequently decompose into the ribosyl-rifamycins initially detected by Dabbs *et al.* The hydroxyl group modified by Arr forms critical hydrogen bonds with RNAP (**Figure 8**); losing this essential interaction and adding a bulky group abolishes the antibacterial activity of rifamycins. Arr is unique among ADP-ribosyltransferases (ARTs) because of its compact size (19kDa) and is the only known ART that modifies a small molecule<sup>192</sup>. ADP-ribosylation of proteins is widespread in eukaryotic organisms, where it plays a role in signalling pathways<sup>193</sup>. In prokaryotes, pathogens widely use ARTs as toxins that inactivate host defences through ADP-ribosylation. The cholera toxin, for instance, ADP-ribosylates host intestinal Gs $\alpha$  protein to cause diarrhea<sup>193</sup>. ARTs are common cargo of type VI secretion systems where they act on DNA, RNA, and protein targets to kill nearby



**Figure 10 The rifamycin ADP-ribosyltransferase Arr.** The hydroxyl linked to the C23 (O23) attacks an oxocarbenium intermediate of NAD to form ADP-ribosyl-rifamycins. The spontaneous loss of ADP which yields the ribosylated species that accumulate in culture media is also depicted.

cells<sup>194,195</sup>. Evolution has effectively repurposed an ancient ART into a resistance enzyme, a testament to the selective pressure of rifamycin producers on their neighbours.

Many mycobacteria are intrinsically rifamycin-resistant because they possess Arr enzymes. The most notorious is *Mycobacterium abscessus* which can cause disease in patients with chronic lung diseases and is also becoming an increasingly common nosocomial pathogen. It is a ubiquitous organism in soil and water that displays intrinsic resistance to many anti-mycobacterial agents, making it very challenging to treat<sup>196–198</sup>. Counterintuitively, *arr* has been mobilized into many Gram-negative pathogens even though rifamycins are not used to treat them. An integron containing *arr* was found in *P. aeruginosa* in 1999 and has since spread widely among  $\gamma$ -Proteobacteria<sup>199</sup>. Routine susceptibility testing panels for Gram-negative pathogens don't include rifamycins; therefore, whole genome sequencing better approximates their prevalence. According to the Comprehensive Antibiotic Resistance Database (CARD), sequenced genomes of 1.74% of *Acinetobacter baumannii*, 0.49% of *P. aeruginosa*, 0.63% of *E. coli*, and 3.06% of *Klebsiella pneumoniae* contain *arr*. In some less common pathogens, *arr* is alarmingly frequent; for instance, 9.09%, 14.71%, and 9.86% of *Providencia stuartii*, *Providencia rettgeri*, and *Escherichia fergusonii* genomes, respectively<sup>200,201</sup>. Dissemination of these resistance elements in off-target bacteria could be a byproduct of the quantity and duration of rifamycin use during tuberculosis therapy. Although *arr-2* mobilization predates the approval of rifaximin, its use in gastrointestinal disorders could also be driving the dissemination of *arr-2*, especially within Enterobacteriaceae<sup>108,202</sup>. Most of the time, the mobilization of a resistance gene into pathogens is cause for widespread concern; Arr has

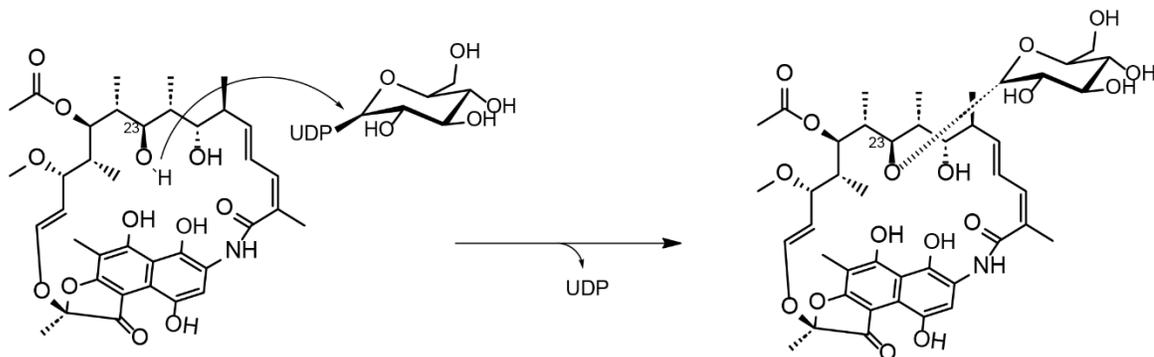
effectively accomplished this jump without much attention. Using membrane-disrupting compounds to increase the penetration of rifamycins into Gram-negative pathogens has received significant interest as a possible therapy; Arr has the potential to impede this promising avenue of research.

The presence of Arr in many mycobacteria has led to considerable interest in developing rifamycin derivatives that can evade this enzyme. Adding bulky moieties to the acetyl group on O25, adjacent to the target of Arr, O23, has been shown on multiple occasions to evade Arr-mediated resistance<sup>203,204</sup>. Most promising is the combination of these O25 modifications with the rifabutin scaffold, which boasts nanomolar MICs against *M. abscessus*<sup>205</sup>. Additionally, Kanglemycin A is a poor substrate for *M. smegmatis* Arr but unfortunately does not escape modification by the *M. abscessus* enzyme making its utility for this purpose questionable<sup>206</sup>. These studies demonstrate that once resistance mechanisms are identified and understood, they can often be overcome by traditional medicinal chemistry approaches and underscore the importance of characterizing all resistance mechanisms that exist in nature.

## **RGT**

Glycosylation of rifamycins has been observed primarily in *Nocardia* and *Streptomyces spp.*<sup>207,208</sup>. First discovered in the 1990s, it was only in 2014 that the gene responsible for glycosylation, termed *rgt*, was finally characterized<sup>209</sup>. Rgt uses an abundant activated sugar, UDP-glucose, as a substrate to glucosylate the O23 residue of rifamycins (**Figure 11**). The same essential hydroxyl is ribosylated by Arr, suggesting it interferes with RNAP binding in the same manner. No structure is available for Rgt, so

mechanistic details on this enzyme are sparse. Based on sequence homology, Rgt is related to biosynthetic glycosyltransferases that decorate antibiotics and other small molecules in actinomycetes. Rgt is most closely related to glycosyltransferases used to synthesize the glycopeptide antibiotics such as vancomycin<sup>209</sup>. Within the actinomycete collection screened in our lab, glycosylation was the most prominent mechanism of inactivation<sup>208</sup>.



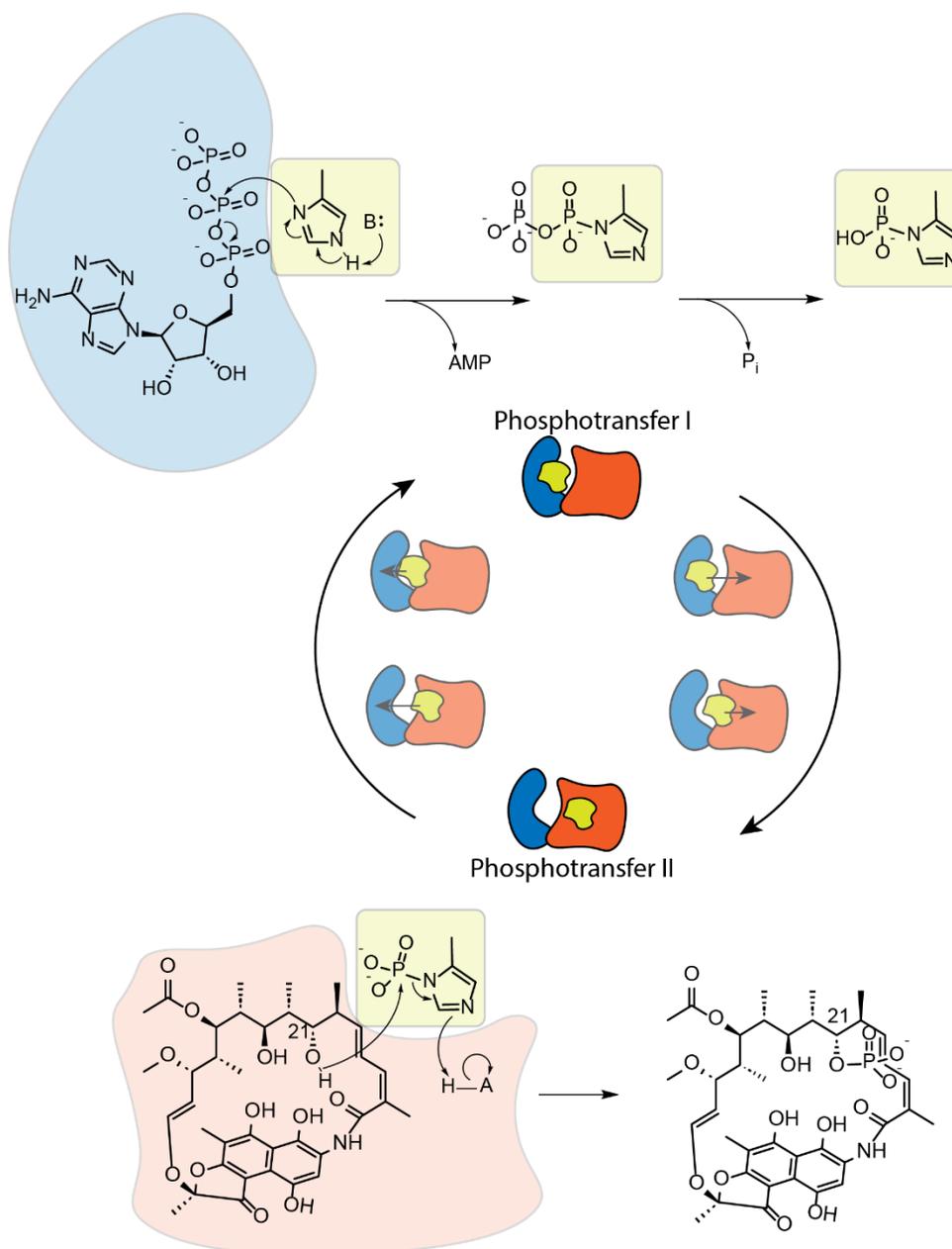
**Figure 11 The rifamycin glycosyltransferase Rgt.** This enzyme uses UDP-glucose as a substrate to glycosylate the O23 hydroxyl of rifamycins.

## RPH

Unlike Arr and Rgt, the rifamycin phosphotransferase (Rph) does not target the O23 hydroxyl and instead phosphorylates the adjacent hydroxyl at position O21(**Figure 12**)<sup>210,211</sup>. This moiety is essential for RNAP inhibition because of its hydrogen bond with RpoB H451, and modifying it results in inactive rifamycin species. This route to inactivation is prevalent in both Firmicutes (*Bacillus* and *Paenibacillus*) and Actinobacteria (*Nocardia* and *Streptomyces*), which can phosphorylate rifamycins<sup>77,208,210,212</sup>. Intriguingly, many rifampin-susceptible organisms, such as *Bacillus cereus*, *Clostridium botulinum*, and *Listeria monocytogenes* encode an Rph. In all cases, these orthologs confer high-level resistance when produced in *E. coli*, so the susceptibility of these bacteria to rifamycins is most likely a consequence of insufficient expression of

*rph*<sup>213</sup>. Future attempts to treat these microbes with rifamycins may inadvertently activate these silent resistance mechanisms.

Rph is unique in its chemical and enzymatic mechanism among the antibiotic-inactivating phosphotransferases. In the case of aminoglycoside and macrolide phosphotransferases (APH and MPH, respectively), these enzymes share homology with protein kinases. In contrast, chloramphenicol phosphotransferase (CMPH) is more closely related to small molecule kinases such as adenylate kinase<sup>211</sup>. APH, MPH, and CMPH enzymes all transfer the  $\gamma$ -phosphate of ATP/GTP to their antibiotic target, releasing ADP/GDP in the process. Rph is not related to classical protein or small molecule kinases but to core metabolic dikinases such as phosphoenol pyruvate (PEP) synthase<sup>214,215</sup>. Rph enzymes instead generate a phosphohistidine intermediate, resulting from the transfer of both the  $\beta$  and  $\gamma$  phosphates from ATP, followed by a rapid loss of the  $\gamma$ -phosphate. A second phosphorylation reaction transfers what was originally the  $\beta$ -phosphate from Rph to the O21 on rifamycins (**Figure 12**). Large conformation rearrangements facilitate this two-step reaction in Rph, where a small mobile domain shuttles between an ATP binding domain where it becomes phosphorylated and a rifamycin binding domain where it phosphorylates its target<sup>214,215</sup>. Bacteria have access to remarkable genetic and enzymatic diversity that can be harnessed by evolution to evade and inactivate harmful molecules like rifamycins. A logical *a priori* assumption would be that a rifamycin phosphorylating enzyme would be a relatively typical kinase. Rph is an example of how evolution can subvert these expectations by co-opting a complex core metabolic enzyme for antibiotic resistance.



**Figure 12 The rifamycin phosphotransferase Rph.** Rph enzymes perform two distinct phosphorylation reactions. The first uses ATP to phosphorylate a histidine residue in Rph, and the second transfers this phosphate to O21 on rifamycins. Dynamic movement of the phospho-histidine swivel domain (yellow) between the ATP binding domain (blue) and the rifamycin binding domain (orange) required for activity are depicted.

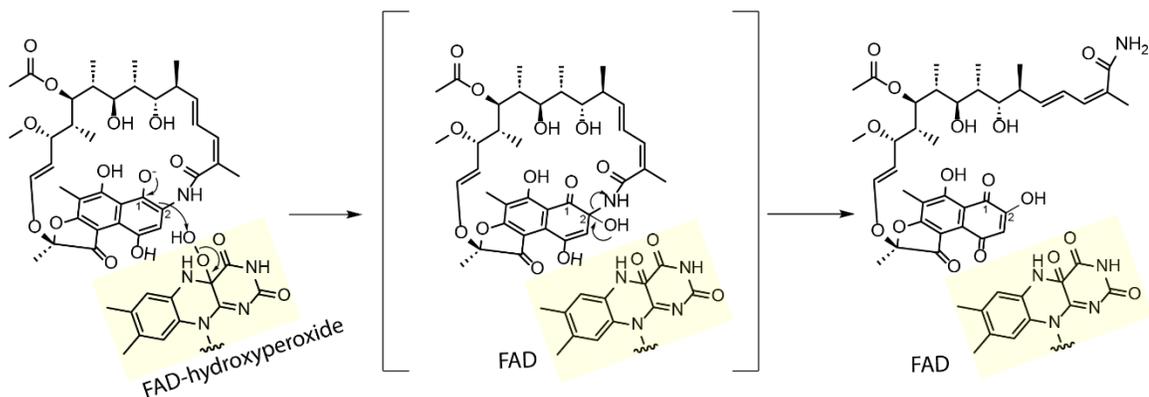
## **Rox/Iri**

In the case of Rph, Arr, and Rgt, elucidating the structure of inactivated products was relatively straightforward. The characteristic orange-red colour of rifamycins aided researchers by simplifying the identification and purification of modified species. However, researchers often identified microbes that could ‘decolorize’ or ‘decompose’ rifampin and thus could not identify the reaction product(s). In 1997 the gene responsible for this phenotype, termed *iri*, was first identified in *Rhodococcus equi*<sup>216</sup>. Iri belongs to the Class A FAD monooxygenases; typically, the FAD cofactor becomes reduced by NADPH and then reacts with O<sub>2</sub> to generate reactive flavin species that can hydroxylate a wide range of substrates<sup>217</sup>. Despite knowing that Iri must inactivate rifamycins through hydroxylation, the modification made to rifamycins was enigmatic for many years. In 2010 Hoshino *et al.* successfully purified the reaction product of an Iri homolog, Rox, from *Nocardia farcinia*. According to their structural assignment, rifampin became hydroxylated on the semi-synthetic tail<sup>218</sup>. This modification was hard to rationalize for several reasons. Why this product would be inactive isn’t apparent because this region does not interact with RNAP. Furthermore, Iri/Rox enzymes can inactivate rifamycins which lack an amine at an equivalent position. Lastly, the hallmark of these enzymes is their ability to ‘decolorize’ rifamycins; the modification proposed by Hoshino *et al.* leaves the primary chromophore of rifampin intact. Working with a homolog from *S. venezuelae*, we determined that Iri/Rox enzymes instead hydroxylate the C2 position of the naphthoquinone core, resulting in C-N bond cleavage that effectively linearizes rifamycins (**Figure 13**)<sup>219,220</sup>. Iri/Rox enzymes open the rifamycin macrocycle, destroying the precise

three-dimensional shape required to engage RpoB and inhibit transcription. Lastly, Rox modifies the chromophore that gives rifamycins their colour, reconciling the initial decolorization observations. Many isolates in our library appear to further degrade rifamycins past the linearized form that accumulates in *N. farcinia* and *S. venezuelae*<sup>208,219</sup>. The transformations that occur post-Iri/Rox are still unknown, but since the modified species generated by Rox are inactive, any downstream modifications are unlikely to have a role in resistance. Intriguingly, the linearization of rifamycins by Rox is an essential step in the biosynthesis of saliniketals by *Salinispora arenicola*. Saliniketals are derived from a rifamycin SV precursor, rifamycin W, cleaved from the aromatic core at both ends of the ansa chain by the combined action of a Rox enzyme and a cytochrome P450 during biosynthesis<sup>221,222</sup>.

Class A FAD-dependent monooxygenases act on a wide variety of substrates. Their specificity towards not only rifamycins but the C2 position requires the correct spatial positioning of the reactive flavin intermediate and the formation of a phenolate anion at the O1 position of rifamycins. The phenolate anion allows for nucleophilic attack of the hydroperoxide flavin from the C2 position (**Figure 13**)<sup>219</sup>. This reaction creates an unstable tetrahedral intermediate at C2 that favours the subsequent breaking of the C-N bond. Natural product rifamycins almost universally contain the hydroxyl at the C1 position, but as Sensi determined, this position can be oxidized to a ketone without compromising bioactivity<sup>82</sup>. Consequently, rifamycins with ketones at this position remain active while being inherently immune to Rox enzymes. The oxidized form of rifamycin SV, rifamycin S, cannot be modified by Rox and even functions as a competitive inhibitor *in vitro*<sup>219</sup>.

Fixed ketones at the C1 position, like the one found in rifabutin, should be incorporated into next-generation rifamycins to evade Rox enzymes.

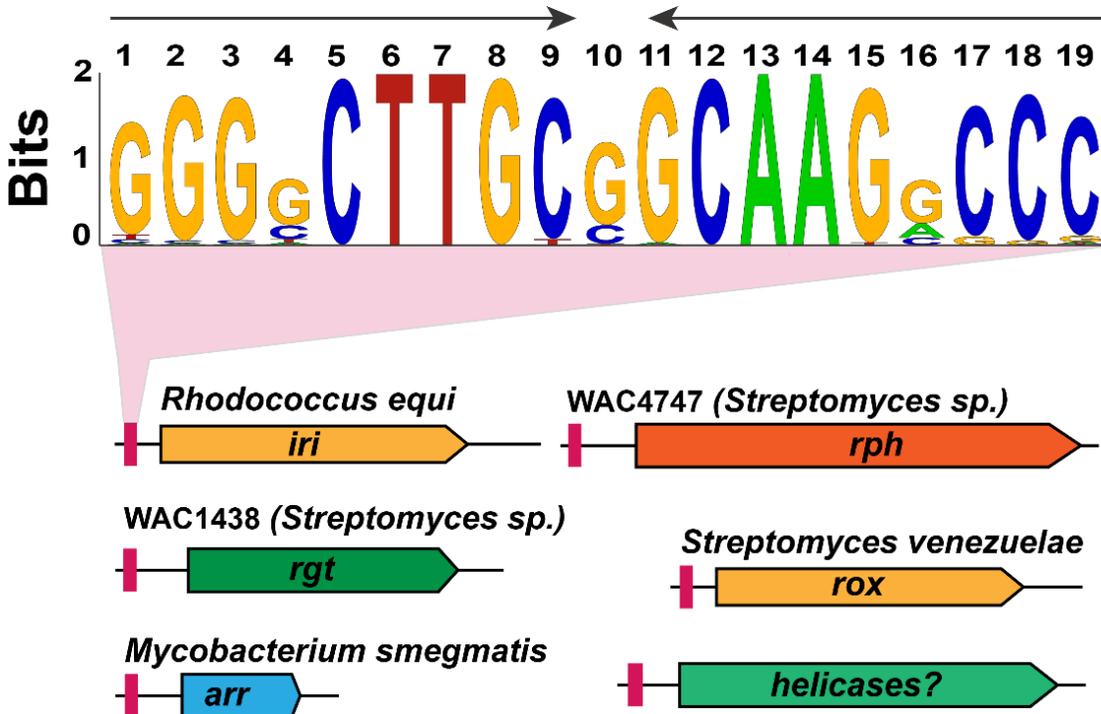


**Figure 13** The rifamycin monooxygenase Rox/Iri. A phenolate intermediate at C1 enables nucleophilic attack from C2 to capture a hydroxyl group from the reactive flavin species C4a hydroperoxyflavin adenine dinucleotide (FAD-hydroperoxide). C-N bond cleavage detoxifies rifamycins by opening their macrocycle.

### Inducible rifamycin resistance and the RAE

The widespread mobilization of dedicated resistance enzymes like  $\beta$ -lactamases has primarily fueled the antibiotic resistance crisis. In this context, these proteins are usually expressed constitutively, even in the absence of antibiotic<sup>223</sup>. Intrinsic resistance mechanisms have been fixed in the genome for millions of years; as a result, the expression of these genes has become tightly regulated. It has been known since the late 1980s that the production of rifamycin inactivating enzymes is inducible by rifamycin antibiotics<sup>211</sup>. More recently, an examination of the DNA upstream of rifamycin-inactivating enzymes revealed the presence of a highly conserved 19bp palindromic sequence, the rifamycin-associated element (RAE) (**Figure 14**). Subsequent experiments showed that this sequence was necessary for the expression of *rph* in response to sub-MIC rifamycin antibiotics. All other antibiotic classes tested failed to induce expression. The RAE, therefore, functions

as a *cis*-regulatory genetic element that controls the expression of downstream genes, specifically in response to rifamycins<sup>213</sup>. However, the mechanism by which rifamycin-mediated induction occurs, and the role of the RAE, are unknown.



**Figure 14 The rifamycin-associated element (RAE).** The RAE is associated with all known rifamycin-inactivating enzymes and putative novel resistance genes. Consensus RAE sequence is depicted as a sequence logo.

The RAE is associated with nearly all instances of *arr*, *rox*, *rph*, and *rgt* in Actinobacteria. Although many Firmicutes can inactivate rifamycins and possess homologs of *rph* and *rox*, this sequence is absent from this phyla<sup>208,213</sup>. The ubiquity of this association raises interesting questions about the evolution and ‘domestication’ of resistance genes, whereby they become tightly regulated over time. RAE-associated genes are highly likely to be involved in rifamycin resistance, meaning the RAE can guide our

goal to characterize the rifamycin resistome. Indeed, preliminary queries of sequence databases show an association of the RAE with a family of uncharacterized proteins annotated as putative helicases, which we hypothesize may represent a novel mechanism of rifamycin resistance.

## **AIMS OF THIS WORK**

As we enter the resistance era, studying the larger antibiotic resistome, including genes and mechanisms not found in pathogens, can bolster our ability to anticipate and respond to emerging forms of antibiotic resistance. The work presented here aims to expand our understanding of the rifamycin resistome and to discover novel RNAP inhibitors, including rifamycins which can evade existing resistance mechanisms to serve as a scaffold for future drug development.

First, I set out to understand the molecular mechanism of induction through the RAE (**Chapter 2**). While ultimately unsuccessful, the work presented here significantly shifts our understanding of this phenomenon. First, we identify additional conserved sequences and spacing requirements important for RAE function. Next, in direct conflict with prior work, we establish that induction requires RNAP inhibition by rifamycins. Although questions remain about the precise mechanism of induction, these observations would set the stage for the work presented in **Chapters 3** and **5**.

Palindromic sequences like the RAE are directionless. In **Chapter 2** we characterize an extended RAE, which includes core promoter sequences and precise spacing requirements. In **Chapter 3** we exploit this to catalogue all RAE-associated genes in Actinobacterial genomes. The RAE, and by extension, rifamycin resistance genes, are

far more abundant than previously realized. Putative helicases (HelRs) are the most abundant genes and likely the last major uncharacterized protein family. Clear taxonomic biases for RAE-associated genes exist alongside widespread functional redundancy (multiple resistance mechanisms) and a bias towards HelR-inactivating gene combinations.

**Chapter 4** consists of our biochemical and *in-vitro* investigation of the RAE-associated ‘helicases’. These proteins are not true helicases and instead bind to RNAP and displace rifamycins to confer broad-spectrum rifamycin resistance. These remarkable enzymes represent a novel and widespread mechanism of rifamycin resistance. They are also the first description of a protection protein that confers antibiotic resistance to an RNAP-targeting antimicrobial.

With the revelation that induction through the RAE occurs following inhibition of RNAP, we used the RAE as a reporter to find novel RNAP inhibitors in our in-house collection of >10 000 microbial extracts. In **Chapter 5** we discover almost 20 new rifamycin producers from our library and identified the rare rifamycin congeners chaxamycin D and sporalactam B, the latter of which we demonstrate can evade the rifamycin phosphotransferase Rph. Additionally, whole genome sequencing of our hits identified multiple biosynthetic gene clusters, seemingly encoding for halogenated rifamycins. I successfully cloned and activated the expression of this ~85kbp cluster in *Streptomyces coelicolor*, setting the stage for future investigation of these uncharacterized rifamycins.

Lastly, the implications and future directions arising from this work are presented in **Chapter 6**.

**CHAPTER 2**  
**Induction through the RAE**

## **PREFACE**

The work presented in Chapter 2 is a manuscript in preparation for submission.

Surette, M.D. and Wright, G.D.

MDS designed and performed the experiments, performed analysis, made figures, and edited and wrote the manuscript. GDW designed experiments, wrote and edited the manuscript

## ABSTRACT

Rifamycin (Rif) antibiotics are a precious clinical resource due to their role in tuberculosis therapy. These compounds can selectively inhibit prokaryotic DNA-dependent RNA-polymerase, giving them broad-spectrum antibiotic activity. Many actinomycete bacteria, which are soil saprophytes and prolific producers of antibiotics, encode inducible Rif resistance. Four distinct enzyme families that can inactivate Rif have been identified, phosphotransferases, glycosyltransferases, ADP-ribosyltransferases, and monooxygenases. A highly conserved 19bp inverted repeat sequence called the Rif associated element (RAE) coordinates the induction of these genes through an unknown mechanism in response to rifamycin antibiotics. Using *Streptomyces venezuelae* as a model, we identify an extended RAE motif that includes core promoter elements and show that it has specific spacing requirements for proper induction. Next, in contrast to previous results, we establish that induction of the RAE requires that Rif interact with their binding site on the  $\beta$ -subunit. We show that *S. venezuelae* mutants with Rif insensitive RNAP are defective in induction. From a panel of RNAP inhibitors, we identified sorangicin as the only non-Rif capable of inducing the RAE. Sorangicin shares a binding site with Rifs, and induction of the RAE by Sorangicin is also perturbed in cells expressing Rif-resistant RNAP. Lastly, we show that a RAE-containing promoter,  $P_{rox}$ , is not intrinsically Rif-resistant *in vitro*. The RAE is able to sense context-specific inhibition of RNAP by Rifs and sorangicin through an unusual but highly conserved DNA element. Although this work falls short of elucidating mechanism, it significantly advances our understanding of inducible rifamycin resistance.

## INTRODUCTION

Antibiotic resistance is a global crisis that currently takes the lives of millions each year, and forecasts suggest this trend will continue to worsen in the coming decades<sup>1-3</sup>. Previously treatable infections have become untreatable by acquiring foreign resistance genes through horizontal gene transfer. Mobilized genes, such as those found on plasmids, transposons, and integrons, are typically expressed continuously. Human pathogens are under constant selective pressure to evade antibiotics. The constitutive nature of many mobilized resistance genes facilitates their dissemination as they can provide protection without requiring or interacting with additional regulatory factors<sup>4-6</sup>. Despite being a modern problem, antibiotic resistance is an ancient trait. Bacteria with intrinsic resistance mechanisms have likely harboured them for millennia and have developed strategies to control their expression, producing them only when necessary<sup>7</sup>. Managing the production of resistance genes is especially important when they carry a fitness cost, such as *van* operons which modify the host's peptidoglycan, or the energetically costly TetA efflux pump<sup>8,9</sup>. It also allows for the high-level production of resistance genes in the presence of an antibiotic, where fitness might be compromised at this expression level in the absence of antibiotics. Therefore, a holistic understanding of antibiotic resistance must include the regulatory mechanisms bacteria use to express resistance genes.

Rifamycins are natural product antibiotics that inhibit the production of full-length RNAs from bacterial DNA-dependent RNA polymerases (RNAP)<sup>10</sup>. These compounds bind to the  $\beta$ -subunit of RNAP and impede the path of nascent transcripts at roughly the

3rd nucleotide position. As a result, rifamycins prevent the synthesis of transcripts longer than 3nt and cause the accumulation of 2-3nt abortive products<sup>11-13</sup>. Since the introduction of rifampin in the late 1960s (a semisynthetic rifamycin derivative), these compounds have been critical in treating *Mycobacterium tuberculosis* infections. The WHO classifies several rifamycins as essential medicines<sup>14</sup>. Unfortunately, resistance arises rapidly to rifamycins due to spontaneous mutations in the  $\beta$ -subunit of RNAP, *rpoB*. Most resistance emerges due to substitutions in a single stretch of 27 amino acids that comprise most of the rifamycin binding pocket. The frequency and magnitude of these mutations manifest in frequent treatment failure during rifamycin monotherapy<sup>15,16</sup>. The use of rifamycins is therefore characterized by a reliance on combination therapies, where rifamycins are used in conjunction with other antibiotics to suppress the development of spontaneous resistance.

From a clinical perspective, the study of rifamycin resistance is essentially synonymous with the study of *rpoB* mutations. In contrast, environmental microbes, particularly those inhabiting the soil, encode a diverse array of rifamycin resistance genes<sup>17</sup>. Enzymes that inactivate rifamycins by ADP-ribosylation (Arr), phosphorylation (Rph), glycosylation (Rgt), and hydroxylation (Iri/Rox) have all been characterized in actinomycetes<sup>18-23</sup>. Even more remarkable is the association of all these genes with a highly conserved 19bp palindromic DNA sequence, the rifamycin-associated element (RAE)<sup>20</sup>.

The RAE is a broadly conserved *cis*-regulatory sequence required to orchestrate the production of rifamycin resistance genes in response to rifamycin antibiotics<sup>20,24</sup>. The RAE is 19bp in length, containing two 9bp inverted repeats with a 1bp spacer. This motif

displays perfect or nearly perfect palindromicity and a high degree of conservation at the sequence level. Expression of genes downstream from a RAE is increased following exposure to sub-inhibitory rifamycins. Furthermore, the RAE is necessary for this to occur, as scrambling this sequence abolishes induction. Reporter constructs using the DNA upstream from an *rph* from *Streptomyces* sp. WAC4747, which contains a RAE, showed that induction is specific to the rifamycin class of antibiotics. Antibiotics targeting DNA gyrase/topoisomerase (novobiocin and ciprofloxacin), the cell wall (ampicillin and vancomycin), and the ribosome (erythromycin and streptomycin) all failed to induce the RAE. Using this reporter, Spanogiannopoulos *et al.* also showed that the RAE functioned in a wide variety of *Streptomyces* that lack their own RAE, including rifamycin-susceptible strains and those that possess a low-level rifamycin resistance due to RpoB substitutions like *Streptomyces coelicolor*<sup>24</sup>. Lastly, Spanogiannopoulos *et al.* raised spontaneous high-level rifampin-resistant mutants and assayed them with their reporter. These strains contained well-documented resistance substitutions and showed equivalent RAE induction to the wild type<sup>20</sup>. These results suggest a widespread and conserved mechanism among actinomycetes to sense rifamycins and respond to them through the RAE.

A major theme in transcriptional regulation is the modulation of RNAP by DNA-binding regulatory proteins which bind and recognize specific DNA sequences. These sequences almost always take the form of direct or inverted repeats<sup>25–27</sup>. Due to these similarities, we hypothesized that the RAE is likely a protein binding site. Next, we reasoned that if a downstream product or side effect resulting from RNAP inhibition was used to sense rifamycins, induction through the RAE shouldn't occur in actinomycetes with

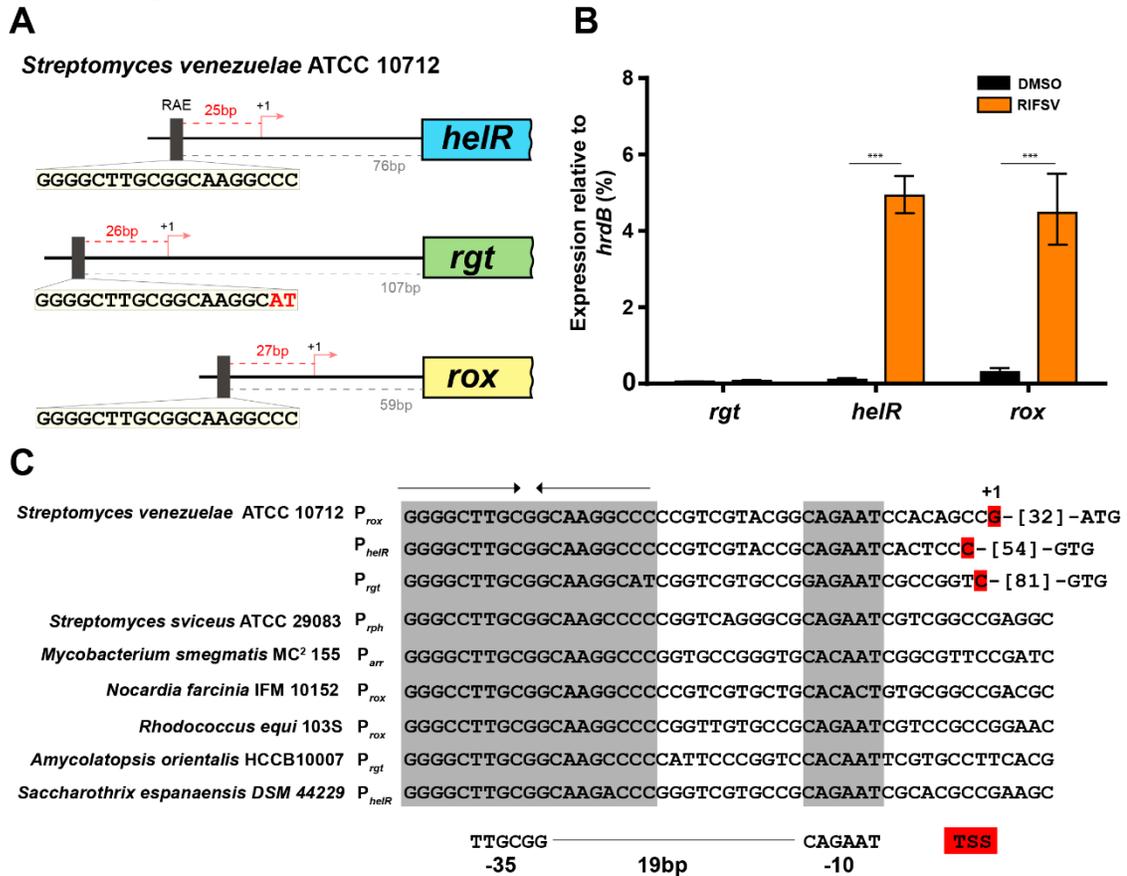
rifamycin-resistant RNAPs. Therefore, it's most likely that rifamycin compounds are sensed directly during induction of the RAE, as this can occur irrespective of RNAP susceptibility or inhibition. From the available data, we hypothesized that the RAE is the binding site for a transcriptional regulator that must interact with rifamycins directly or with a dedicated rifamycin sensing protein, as in the case of two-component systems<sup>26</sup>. Therefore, we undertook the work presented herein to find and characterize this RAE-binding protein.

## RESULTS

To study the RAE, we chose the genetically tractable model organism *Streptomyces venezuelae* ATCC 10712. It was the only model *Streptomyces* that possessed at least one RAE. Previous work demonstrating that the *rox* enzyme in this bacterium is inducible by rifamycins also indicated this would be a sound model system. A query of the *S. venezuelae* genome revealed that it encoded two RAEs controlling *rox* and *SVEN6029*, a putative helicase (*helR*), respectively (**Figure 1A**). In addition, it encodes a third locus highly similar to the RAE but with a 2bp mismatch at the 3' end (CC to AT), upstream of a putative glycosyltransferase (*rgt*)(**Figure 1A**).

We examined the expression of all three RAE-associated genes using RT-qPCR following a 2h treatment with 0.5 µg/mL rifamycin SV or DMSO control. As shown in (**Figure 1B**), both *rox* and *helR* are inducible by sub-inhibitory concentrations of rifamycin antibiotics. Expression of *helR* and *rox* increased by approximately 50 and 15-fold, respectively, following treatment with rifamycin SV. Higher basal transcription of *rox* gives the impression of lower induction, but both transcripts reach a comparable final

expression level. While investigating the rifamycin monooxygenase Rox, we extensively analyzed inactivated rifamycins produced by *S. venezuelae* and never observed any glycosylated species<sup>23,28</sup>. Consistent with this experience,  $P_{rgt}$  is non-inducible by rifamycins, presumably due to its 3' mismatch to the consensus RAE sequence.

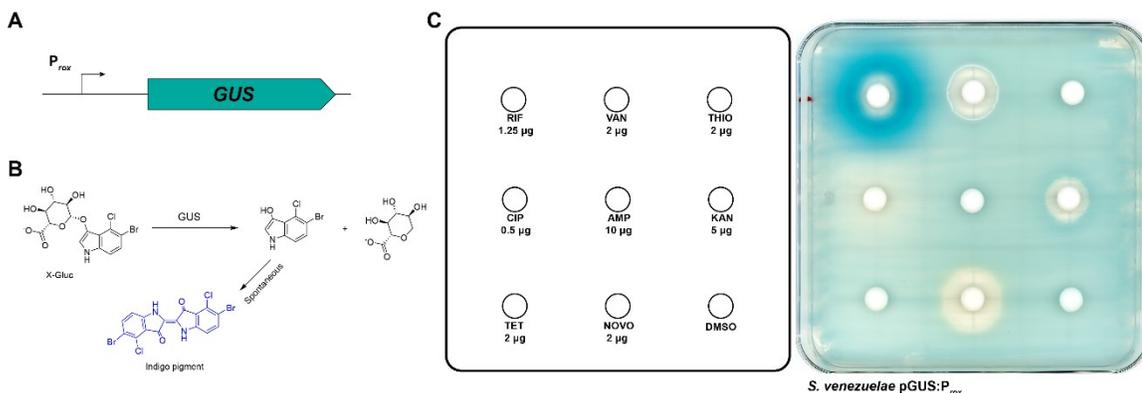


**Figure 1 A) The RAEs in *S. venezuelae* and their transcription start sites. B) *rox* and *helR* are inducible by rifamycins.** Relative expression of *rox*, *rgt*, and *helR* after treatment with sub-inhibitory rifamycin SV or vehicle control (DMSO) were quantified by RT-qPCR. Expression is normalized to the housekeeping gene *hrdB*. Error bars represent standard deviation. \*\*\* $p < 0.001$  according to student's t-test. **C) The extended RAE** Alignment of the RAE from diverse organisms shows the conservation of a proximal putative -10 site spaced precisely downstream from the inverted repeat motif and an equally well conserved putative -35 site within the RAE.

We then leveraged a pre-existing RNAseq dataset from *S. venezuelae* during vegetative growth to assign putative transcription start sites (TSS) for the RAE-associated genes<sup>29</sup>. The distance from the RAE to the start codon of the gene it controls varies significantly; for instance, in *S. venezuelae*, it ranges 59, 76, and 107bp for *rox*, *helR*, and *rgt*, respectively. In contrast, the three putative TSS were all found within 3bp of each other, approximately 26bp downstream of the RAE (**Figure 1A**). Firstly, this would mean that the RAE is not transcribed, and therefore it cannot be part of a riboswitch or attenuator-like regulatory mechanism. Secondly, the close and conserved distance between RAE and TSS suggested that this element overlaps with a promoter. A motif similar to the *Streptomyces*  $\sigma^{hrdB}$  consensus -10 site (CAGAAT vs. TAGAAT) is present 8-10bp upstream of the TSS, consistent with typical -10 to TSS spacing<sup>30,31</sup>.  $\sigma^{hrdB}$  is the *Streptomyces* housekeeping  $\sigma$ -factor, the equivalent of  $\sigma^{70}$  in Proteobacteria and  $\sigma^A$  in mycobacteria<sup>27,32</sup>. Moreover, the 3D structure of RNAP dictates the spacing of -10 and -35 sites. For  $\sigma^{70}$  family  $\sigma$ -factors like  $\sigma^{hrdB}$ , this spacing is  $17\pm 1$ bp, suggesting the -35 lies within the RAE. A high identity match to the  $\sigma^{hrdB}$  -35 is present in the RAE (TTGCGG vs. TTGACA). However, if these sequences represent -35 and -10 sequences, they are sub-optimally spaced at 19bp<sup>32,33</sup>. Analysis of RAEs from multiple genera demonstrates that this 19bp spacing is strictly conserved (**Figure 1C**). These data suggest that the RAE is part of a larger sequence motif, which we've termed the extended RAE.

To probe the activity and function of RAE-containing promoters, we cloned the ~175bp of the intergenic region upstream of *rox* into the reporter plasmid pGUS<sup>34</sup>. Expression from  $P_{rox}$  produces a colorimetric reaction signifying induction (**Figures 2A**

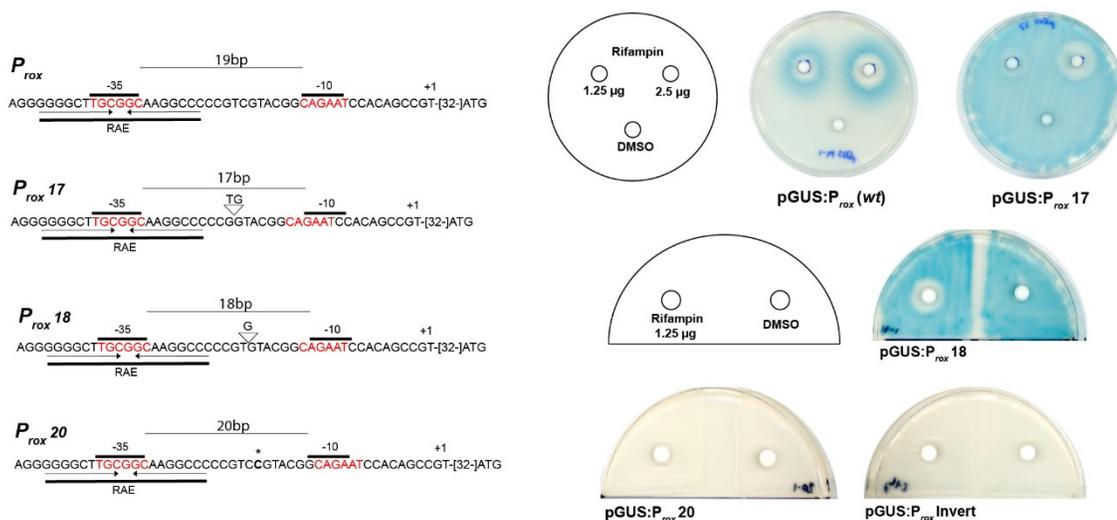
**and B).** We validated that *S. venezuelae* pGUS:P<sub>rox</sub> is specifically induced by rifamycins and not other classes of antibiotics, as shown previously (**Figure 2C**)<sup>20</sup>.



**Figure 2 A reporter system for rifamycin mediated induction through the RAE A) Transcriptional fusion of P<sub>rox</sub> and  $\beta$ -glucuronidase (GUS). B) Production of indigo pigments from X-gluc by GUS. C) Induction of P<sub>rox</sub> is specific to rifamycin antibiotics. *S. venezuelae* was inoculated onto the surface of an agar plate containing X-gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronide) and cellulose discs containing antibiotics were added on top. After incubation for 48 hours, induction is visible as a blue pigment. RIF, rifampin, VAN, vancomycin, THIO, thiostrepton, CIP, ciprofloxacin, AMP, ampicillin, KAN, kanamycin, TET, tetracycline, NOVO, novobiocin.**

If the extended RAE is a promoter with suboptimal spacing, changing this spacing should have predictable effects. Namely, if the distance between the -10 and -35 is shortened, the promoter should become more active. We constructed and assayed P<sub>rox</sub> with 17 and 18bp spacers and observed a significant increase in expression from P<sub>rox</sub> (**Figure 3**). However, these promoters had lost the ability to be induced by rifamycins. Conversely, we observe a complete loss of expression from P<sub>rox</sub> when the spacer is extended by a single nucleotide to 20bp, even in the presence of rifamycins (**Figure 3**). Furthermore, expression is lost when we invert the entire P<sub>rox</sub> intergenic region to direct the extended RAE in the opposite direction effectively. While the RAE is a directionless, palindromic sequence, it

is part of an extended motif that includes a promoter with highly conserved yet suboptimal spacing required for induction.



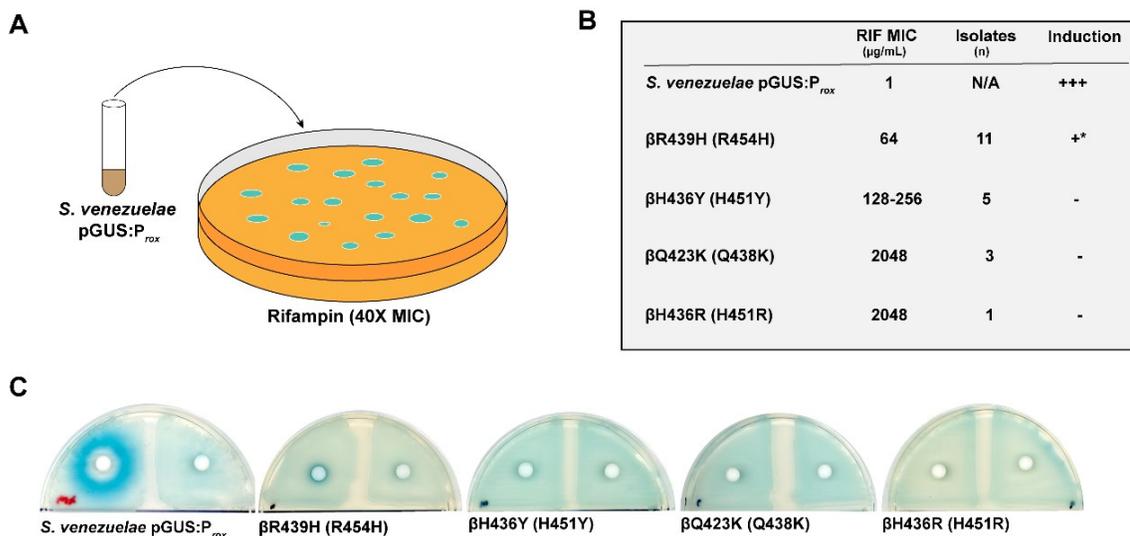
**Figure 3 Conserved spacing of extended RAE elements is required for induction.** Mutations in the spacer region are shown on the left, deletions/insertions were made outside the inverted repeat region. The induction phenotype for each mutant is shown on the right. pGUS: $P_{rox}$ Invert contains the entire wildtype region in the reverse orientation.

These data, taken alongside previous work, lead us to believe that the RAE was the binding site for a transcriptional activator that either sensed rifamycins directly or participated in a signaling cascade with a protein that does. Specifically, we hypothesized that a MerR family regulator was involved. These regulators bind to the spacer region of promoters with longer than optimal spacing. Typically activated by ligand binding, MerRs contort the promoter DNA into a conformation resembling proper spacing in agreement with both the strictly conserved 19bp spacing and our promoter mutants<sup>35</sup>. Furthermore, the 3' mutations in  $P_{rgt}$ , which is non-inducible, could be rationalized by a loss of MerR activator binding. We invested considerable time into traditional approaches, such as using

DNA baits to 'pull down' and identify a specific RAE interacting protein. Different binding buffers, washes, eluents, and even various bacterial species were tried without success. The proteins identified by these experiments were implausible. For example, one protein we pulled down is an extracellular phosphate binding protein (PstS) that couldn't conceivably interact with genomic DNA sequences like the RAE. We also identified proteins that are unique to *Streptomyces* and therefore cannot explain how the RAE might function in *Mycobacterium* or *Nocardia*. Following the data available at the time, we knew that the sensitivity of RNAP did not affect induction, meaning a molecule capable of sensing rifamycin molecules must be present. To this end, we designed and synthesized a photo-affinity probe to label and identify rifamycin-binding proteins. This probe retained both antibiotic activity and the ability to induce the RAE and, therefore, should have been able to identify the putative rifamycin sensor, but this approach also ultimately failed. Serendipitously, a use for this probe did arise in future work (See chapter 4). Next, we compiled a list of MerR regulators conserved between distantly related species with RAEs, overexpressed them in *S. venezuelae*, and looked for perturbed induction of the RAE. None of these constructs had any appreciable change in induction. Given these persistent failures, we decided to replicate some critical experiments in our new *S. venezuelae* system, starting with the effect of rifamycin-resistant *rpoB* alleles on induction.

We raised spontaneous high-level rifampin-resistant mutants by plating *S. venezuelae* on Bennett's media containing 40 µg/mL rifampin (40X the MIC of *S. venezuelae*). In total 20 resistant isolates were recovered, assayed for induction of  $P_{rox}$ , rifampin resistance, and genotyped by sequencing their RRDR (**Figure 4A and B**).

Despite isolating 20 mutants, only 4 alleles were present, RpoB ( $\beta$ )H436R (H451R *M. tuberculosis* numbering),  $\beta$ H436Y (H451Y),  $\beta$ R439H (R454H), and  $\beta$ Q423K (Q438K). These substitutions have been identified as rifampin-resistant in *E. coli* or *M. tuberculosis*<sup>16,36,37</sup>. All four genotypes (and all 20 isolates) show drastically reduced induction compared to the susceptible parent strain (**Figure 4C**). Inconsistent with previously reported results, these data suggest that RAE induction requires rifamycins to bind RNAP. Curiously, the  $\beta$ R439H shows slight but detectable induction of  $P_{rox}$  around the rifampin-containing disc, while the other 3 alleles show no induction. This phenotype was consistent among all 11 isolates with the  $\beta$ R439H substitution. Isolates with the  $\beta$ R439H substitution were also the most susceptible, at 64  $\mu$ g/mL (**Figure 4B**). In Gram

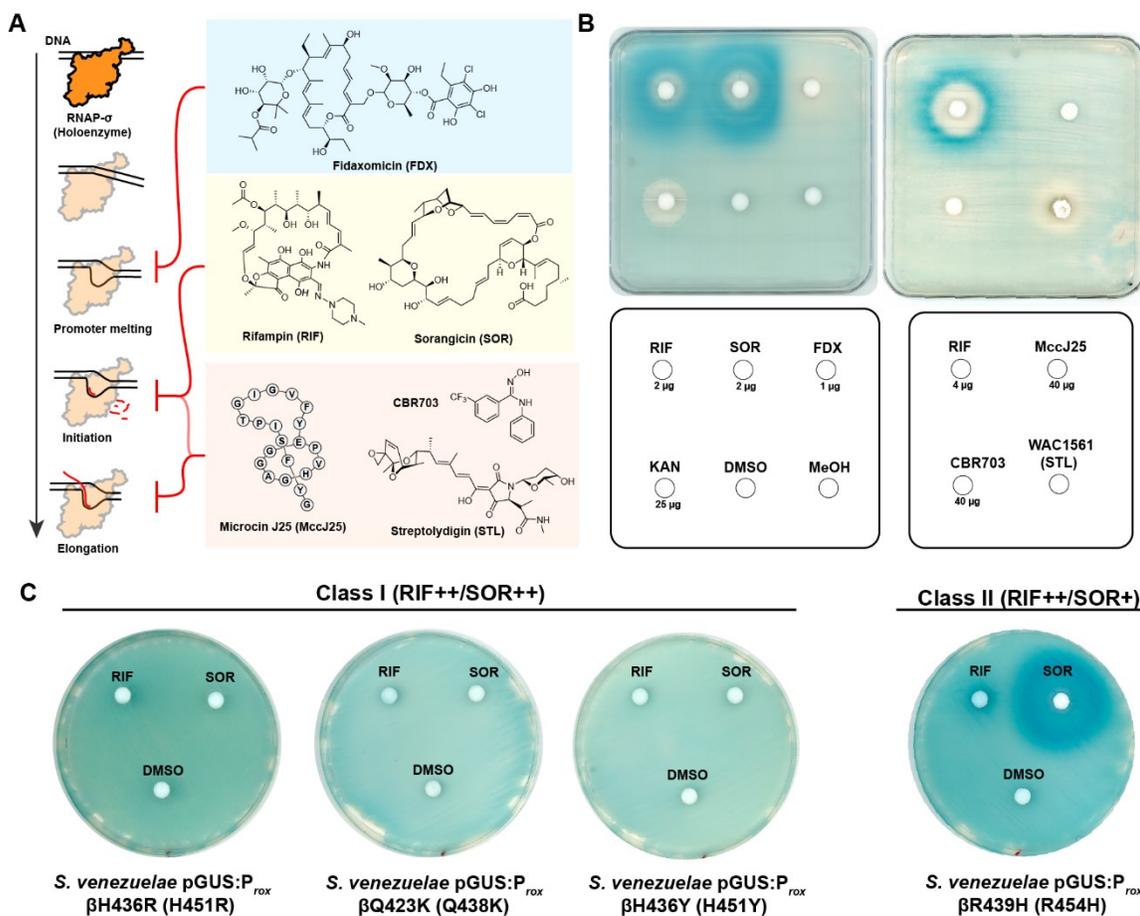


**Figure 4 Rifamycin-resistant RpoB alleles are defective in the induction of the RAE. A) Schematic for isolation of rifamycin-resistant isolates B) Summary table of Rifamycin resistance genotypes and phenotypes C) Induction of *S. venezuelae* harboring rifampin resistant RpoBs. Numbers correspond to *S. venezuelae* RpoB, *M. tuberculosis* number is given in parentheses. +++, wildtype induction, +\* weak induction, - no induction.**

positive organisms, which rarely gain high-level resistance from changes in permeation, the relative differences in MIC are attributable to differences in affinity for RNAP<sup>16,38</sup>. Therefore, RAE induction depends on the ability of rifamycins to bind the  $\beta$ -subunit of RNAP. Furthermore, the behaviour of  $\beta$ R439H suggests a direct inverse correlation between RNAP susceptibility and induction.

We know that antibiotics with different molecular targets than rifamycins (ribosome, topoisomerase, cell wall, etc.) do not induce the RAE<sup>20</sup>. Now that the importance of rifamycin-RNAP interactions in the induction of the RAE is better understood, we were curious if induction is specific to RNAP-acting antibiotics or just the rifamycins. We assayed an assortment of RNAP inhibitors that act through various mechanisms (**Figure 5A**). Sorangicin is a marvel of convergent evolution, a myxobacterial natural product which, despite a highly divergent chemical structure, binds to the same pocket on the  $\beta$ -subunit of RNA polymerase as rifampin<sup>39</sup>. These molecules share a mechanism of action (steric occlusion of the nascent transcript) and even show partial cross-resistance. Fidaxomicin, an actinomycete natural product, acts on an earlier step in the transcription cycle; this antibiotic prevents the conversion of the RNA polymerase-promoter complex from its closed to open conformation, preventing promoter melting<sup>40,41</sup>. We also included a series of later-acting inhibitors CBR703 (Enamine), Microcin J25 (MccJ25), and Streptolydigin (STL) which hinder mobile domains within the active site of RNAP and shut down the nucleotide addition cycle<sup>42,43,44,45</sup>. Unfortunately, MccJ25 and STL are not commercially available. We were able to produce MccJ25 in *E. coli* using an existing expression construct and purify it by reverse-phase chromatography<sup>46,47</sup>. Next, we

searched for the STL biosynthetic gene cluster in genomes from our in-house actinomycete collection and found one in WAC1561. Subsequently, we purified STL from fermentations of WAC1561 (**Supplementary Figure 1**)<sup>48</sup>. Only rifamycins and sorangicin are inducers



**Figure 5 Induction of the RAE is specific to antibiotics targeting the rifamycin binding pocket. A) RNAP inhibitors and where they act in the transcription cycle. B) Induction of the RAE by RNAP inhibitors. RIF, rifampin, SOR, sorangicin, FDX, fidaxomicin, KAN, kanamycin, MccJ25, microcin J25, STL, streptolydigin, MeOH, methanol. C) Differential susceptibility to sorangicin correlates with RAE induction. RIF++/SOR++, indicates roughly equivalent high-level resistance to rifampin and sorangicin. RIF++/SOR+ indicates high-level resistance to rifampin but moderate resistance to sorangicin.**

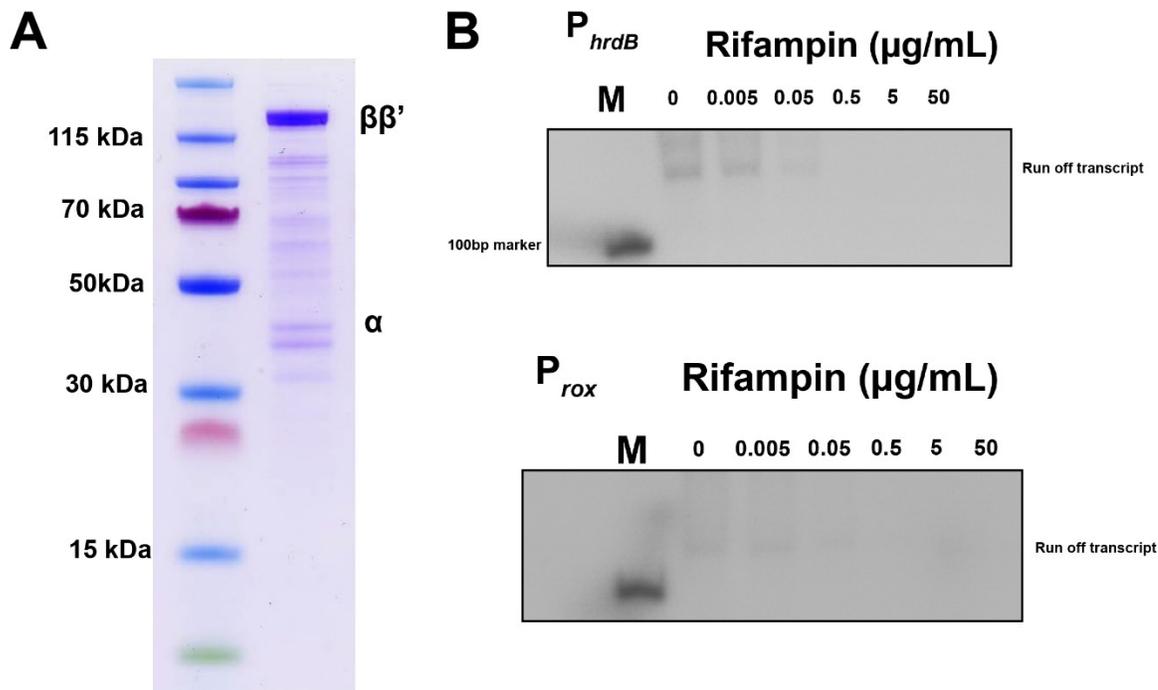
of  $P_{rox}$  (**Figure 5B**). All other compounds either had no antimicrobial activity against *S. venezuelae*, as in the case of CBR703 and MccJ25, or, like FDX and STL, could inhibit *S. venezuelae* but not induce the RAE. We also note the modest antagonistic activity of fidaxomicin towards induction; this effect is reproducible but doesn't translate to significant rifampin-fidaxomicin synergy against *S. venezuelae* (**Supplemental Figure 2**). Induction of the RAE is therefore restricted to molecules that specifically interact with the rifamycin binding pocket on RpoB, such as sorangicin.

While rifamycins and sorangicin share a mechanism of action and binding pocket on RNAP, differences in the specific molecular interactions between sorangicin and RpoB manifest as altered susceptibility to certain RRDR substitutions. Many alleles confer a roughly equivalent level of resistance to rifampin and sorangicin, termed class I substitutions by Campbell *et al.*; class II alleles confer more resistance to rifampin than sorangicin<sup>39</sup>. Of the genotypes we isolated,  $\beta$ R439H (R456H) and  $\beta$ H436Y (H451Y) were shown previously by Campbell *et al.* to belong to class II and I, respectively. We reasoned that if the ability to bind RNAP is the sole determinant for induction by rifamycins and sorangicin, then the class II mutant should show higher susceptibility to sorangicin and, therefore, increased induction of the RAE relative to rifampin. We tested representatives from each genotype against both rifampin and sorangicin (**Figure 4C**). Campbell *et al.* did not previously characterize  $\beta$ H436R and  $\beta$ Q423K sorangicin susceptibility, but based on our results, we have putatively assigned them to class I as they show no induction by sorangicin or rifampin, indicative of roughly equivalent resistance. The previously classified alleles behaved as expected;  $\beta$ H436Y (class I) was non-inducible by both drugs,

whereas  $\beta$ R439H showed faint rifampin induction but significant induction by sorangicin. The ability of rifamycins *and* sorangicin to induce the RAE is dependent on their affinity for RNAP.

We next considered whether rifampin directly influenced transcription from  $P_{rox}$ . If  $P_{rox}$  is an innately rifamycin-resistant promoter, perhaps induction at this locus is a passive process whereby free RNAPs are biased towards initiating transcription here because rifamycin-inhibited RNAPs are blocking other promoters. Another possibility is that there is something unique about the structure of the RNAP- $P_{rox}$  complex, which makes rifamycins stimulatory rather than inhibitory. To test these possibilities, we purified native RNAP from *S. venezuelae* (**Figure 6A**). We pooled all RNAP-containing fractions for these experiments to not exclude specific factors or holoenzymes. The  $P_{rox}$  template was amplified from the *S. venezuelae* genome and contained the entire intergenic region upstream of *rox* and 145bp downstream of the predicted TSS (110bp into the ORF). For a control promoter, we amplified a 375bp fragment containing the 275bp preceding the *hrdB* ORF and its first 100bp ( $P_{hrdB}$ ). Both promoters were used as templates in a multi-round run-off transcription assay with *S. venezuelae* RNA polymerase over a gradient of rifampin concentrations (**Figure 6B**). Several transcripts are visible for each template, the appearance of spurious TSS during *in-vitro* transcription is a well-known phenomenon, so it is possible that not all transcripts are produced *in-vivo*. The intense and diffuse signal at the bottom of the gels are short abortive transcripts. Both promoters show significant inhibition of all transcripts beginning at 0.05  $\mu$ g/mL rifampin with complete inhibition at 0.5  $\mu$ g/mL, consistent with previous  $IC_{50}$  measurements from other Actinobacteria<sup>12,49</sup>. The

addition of rifampin does not stimulate or repress the *rox* promoter or lead to the appearance of a new transcript. Together, these data suggest that induction requires additional factors or cannot occur on a relaxed, linear template.



**Figure 6**  $P_{rox}$  is susceptible to rifampin *in vitro*. **A)** Preparation of native *S. venezuelae* RNAP. Coomassie stained SDS-PAGE gel of *S. venezuelae* RNAP. Subunits  $\beta$ ,  $\beta'$ , and  $\alpha$  are labelled. **B)**  $P_{rox}$  and the housekeeping promoter  $P_{hrdB}$  are equally susceptible to rifampin. Autoradiographs of multiple rounds of *in vitro* transcription assays at various rifampin concentrations.

## DISCUSSION

In this study, we sought to understand the mechanism of inducible rifamycin resistance in Actinobacteria and the role of the conserved *cis*-regulatory DNA sequence, the RAE, in this process. While the ultimate role of this element and specific mechanistic insight into rifamycin sensing remains elusive, we have advanced our understanding of this

phenomenon in two crucial ways. First, we have identified additional conserved elements important for induction, including a sub-optimally spaced promoter. Secondly, we have determined that the ability of rifamycins and another class of RNAP inhibitor, sorangicin, to induce expression through the RAE depends on their ability to bind and likely inhibit RNAP. This latter result is in stark contrast to previous data, and together these have significant mechanistic implications for inducible rifamycin resistance.

### **The extended RAE**

The RAE encompasses more conserved sequences than the 19bp inverted repeat originally identified in Spanogiannopoulos *et al.* 2014<sup>20</sup>. We identify core promoter motifs (the -35 and the -10), which align with a transcription start site just downstream of the RAE (**Figure 1**). The -35 site lies within the RAE and is spaced exactly 19bp away from a highly conserved -10 site across instances of the RAE found from multiple genera that regulate diverse resistance enzymes. The distance of 19bp is significant because it is too far to be efficiently bound by RNAP. Sequence-specific recognition of promoters by  $\sigma$ -factors requires spacing of 17 +/- 1bp to make the major groove accessible to the holoenzyme at the -35 and -10 sites. Spacers that are too long will pull or push these motifs out of alignment, making transcription inefficient or impossible. The strict conservation of sub-optimal spacing is not a coincidence, as shortening this spacer led to constitutive expression and a loss of specific induction by rifamycins (**Figure 3**). Conversely, a lengthening of just a single base pair produced a completely inactive promoter. The RAE is not just near a promoter; the RAE is effectively part of one. Lastly, these insights also reveal that the RAE has an innate *direction*, meaning it is now possible to build a comprehensive catalogue of

genes regulated by the RAE across all microbial genomes, explored in Chapter 3. Further work should examine the effect of additional point mutations on induction and try to rationalize the inactivity of  $P_{rgt}$ .

Although the RAE overlaps with core promoter sequences, as many canonical DNA binding transcriptional regulators do, this hypothesis has many problems. A typical transcriptional repressor, like TetR or LacI, binds to inverted repeats and occludes RNAP's access to promoters until they bind their ligands, causing them to dissociate<sup>9,50</sup>. Repressor proteins show considerable flexibility in where they bind relative to the -10 and -35; therefore, the absolute conservation of the spacing between RAE and -10 site would be highly unusual<sup>51</sup>. Furthermore, in the case of our spacer mutants, it is unclear why mutations outside the inverted repeat sequence would result in derepression. If the RAE is not a binding site for a repressor protein, perhaps an activator binds here? Transcriptional activators recruit RNAP to promoters and fall into three general categories based on their mechanisms, class I, class II, and the MerR-like regulators. Class I activators interact with the flexible C-terminal domain of the  $\alpha$  subunit of RNAP ( $\alpha$ CTD); as a consequence, these generally bind farther upstream of the -35 (from 20 to >100bp), often binding in multiple sites. This type of regulator is a poor candidate for a RAE binding protein. Class II activators overlap with the 5' end of the -35 and occupy a space called the -41.5 site. These activators interact with the core body of RNAP, not the flexible  $\alpha$ CTD extensions, and must be positioned precisely. The RAE, however, is centered directly on the -35, highly atypical of class II activators<sup>52-57</sup>. Lastly, the MerR family of regulators was an excellent candidate for this regulator. These bind to promoters with sub-optimal spacing, 19-20nt,

and in the presence of their ligand, undergo a conformational change that distorts the promoter DNA to bring the -10 and -35 sites back into proper alignment. Proper spacing of these promoters is essential for their function. Furthermore, as seen in our work, shortening the spacer produces constitutive non-inducible promoters, and insertions abolish transcription entirely<sup>58,59</sup>. We overexpressed roughly a dozen candidate MerR regulators in *S. venezuelae* and saw no change in induction. Again, apparent differences exist between canonical MerR activity and the RAE. For instance, members of this family typically recognize inverted repeats inside the spacer region, whereas the RAE is centered around the -35<sup>35</sup>. It is possible, however, that knockouts of candidate MerRs would produce a more definitive result. All of the above examples of regulatory systems, from TetA to MerR, all share one critical similarity; they autoregulate themselves. In addition to controlling the expression of target or effector genes, they also tightly regulate their own production to reach high levels of expression quickly and promptly turn off transcription when the stimulus is no longer present. As far as we can tell, the RAE is not associated with any transcriptional regulators, so it is not apparent how inducible rifamycin resistance would tackle these classic problems without the feedback supplied by autoregulation<sup>60,61</sup>. These facts, coupled with our lengthy search for RAE binding proteins and rifamycin binding proteins, lead me to speculate that the RAE is not likely to be the binding site for a classical activator or repressor protein.

The hypothesis that a dedicated  $\sigma$ -factor recognizes the RAE is also worth consideration. Firstly, the conserved -10 site in the extended RAE is a match precisely to the motif found in the housekeeping  $\sigma$ -factors such as  $\sigma^{\text{HrdB}}$ ,  $\sigma^{\text{A}}$ , and  $\sigma^{70}$ . The -10 motifs

recognized by other  $\sigma$ -factors are distinct and usually don't resemble the housekeeping consensus sequence. The putative -35 site we observed is also an excellent match for the housekeeping sequence. Moreover, the typical length for such a site is 6bp, whereas the RAE is almost three times that size making it unlikely that the  $\sigma$ -factor can recognize the whole motif<sup>27</sup>. We are also unaware of any examples of sigma factors that bind to inverted repeats.  $\sigma$ -factors outside the housekeeping family sometimes called extracytoplasmic function (ECF) or group 4  $\sigma$ -factors, are another possibility. These represent the minimal protein machinery for recognizing promoters. Holoenzymes containing these  $\sigma$ -factors are less efficient at initiating transcription because they lack regions equivalent regions, which in the  $\sigma^{70}$  family assist in promoter melting and stabilizing the open complex<sup>27</sup>. The ECF  $\sigma$ -factors need to bind to their cognate -10 and -35 sites with high affinity to compensate and, as a result, are much more specific in the sequences they recognize. Todor *et al.* exploited the tendency of ECF  $\sigma$ -factors to regulate their production, coupled with their high identity requirements for their target promoters, to predict their recognition sites for almost 70% of all ECFs found in bacteria<sup>62</sup>. Again, we do not find any regulatory proteins, such as an ECF  $\sigma$ -factor associated with the RAE (see Chapter 3 for more detail). In my view, it is unlikely that an ECF  $\sigma$ -factor could bind and initiate transcription at its own promoter and then an entirely different sequence, the RAE. These constraints on  $\sigma$ -factor form and function suggest that a dedicated RAE-binding  $\sigma$ -factor is unlikely to exist<sup>63</sup>.

Lastly, we explored the possibility that induction requires only RNAP, the extended RAE, and rifamycins without additional factors. We purified RNAP from *S. venezuelae* and showed that rifampin inhibits transcription from  $P_{rox}$  at concentrations comparable to

a control promoter and that sub-inhibitory rifamycins don't stimulate transcription *in-vitro* either (**Figure 6**). These results indicate that something is missing from our *in-vitro* system required for induction; the nature of what is missing is less certain. Ultimately it could be a DNA or RNAP interacting protein, specific DNA topology imparted by supercoiling/methylation, or even a metabolite. More work is required to fully understand inducible rifamycin resistance and the regulatory and biochemical logic underlying the extended RAE's unique structure.

### **Rifamycins interact with RNAP to induce the RAE**

We provide results in this work that establish that rifamycin-mediated induction depends on the ability of rifamycins to interact with their target, RNAP. While some inducible resistance mechanisms do encode machinery to specifically sense antibiotic molecules themselves, such as the VanR system and TetRA systems, there is also precedent for systems that rely on the molecular mechanism of the antibiotics they sense<sup>64,65</sup>. For instance, AmpR represses the transcription of AmpC  $\beta$ -lactamases until it binds peptidoglycan fragments that accumulate due to  $\beta$ -lactam activity, which causes it to dissociate from the promoter and allow the expression of *ampC*. The elegant leader-peptide attenuator-based mechanisms that control erythromycin resistance, which requires on-ribosome stalling by erythromycin, is another example<sup>66,67</sup>. In both cases, bacteria have exploited precise biochemical cues of antibiotic activity by integrating them into regulatory pathways. Our work suggests that rifamycin-mediated RAE induction follows this logic as well.

Here we show that the ability of rifamycins to bind and presumably inhibit RNAP is required to induce the RAE. For instance, *S. venezuelae* harboring rifamycin-resistant RpoBs show impaired induction (**Figure 4**). Furthermore, we discovered that sorangicin is an equally potent inducer of the RAE. Despite testing RNAP inhibitors with diverse structures, modes of action, and binding sites, only sorangicin, the only other compound known to bind the rifamycin binding pocket, was able to induce the RAE. The unique interaction between rifamycins/sorangicin and their binding pocket inside RNAP is essential for induction.

All available evidence suggests that induction occurs when rifamycins significantly inhibit RNAP. This phenomenon manifests as induction at concentrations near but below the MIC or a blue halo around the zone of inhibition using pGUS: $P_{rox}$ . By making *S. venezuelae* more resistant to rifamycins with rifamycin-resistant RpoB alleles, we've raised the MIC, and accordingly, the sub-MIC induction zone now lies at a higher concentration. For instance, the only allele which retained any hint of induction was  $\beta$ R439H, and susceptibility testing confirmed that this strain was also the most susceptible to rifampin (**Figure 3**). Furthermore,  $\beta$ R439H possesses a high level of rifamycin resistance but only modest sorangicin resistance. As expected, sorangicin produces significant induction from the RAE in this background, but none of the others tested (**Figure 4**)<sup>39</sup>. Conversely, if raising the rifamycin MIC increases the concentrations required for induction, then lowering the rifamycins MIC should result in induction at lower concentrations. We observe this effect in *S. venezuelae*  $\Delta_{rox}$ ,  $\Delta_{helR}$ , and  $\Delta_{rox}\Delta_{helR}$  mutants constructed in Chapter 4, whereby these strains become increasingly susceptible to rifampin and

correspondingly show GUS activity at increasingly lower quantities of rifampin. Induction of the RAE most likely depends on biochemical cues which arise from significant RNAP inhibition. The nature of this cue or cues remains unknown, but it is clearly specific to the activity of rifamycins/sorangicin and not to the inhibition of transcription generally (**Figure 4**). The connection of these cues to the structure/function of RAE loci is also unclear.

### **Future directions**

Inducible rifamycin resistance remains an enigmatic phenomenon, but the work here sets the stage for several avenues of future research. I will briefly highlight those which I consider most promising.

Recent investigations of transcriptional regulation often rely on RNAseq and other omics-based approaches. RNAseq and proteomics experiments have been performed on Mycobacteria exposed to rifampin, including some species which harbor RAEs<sup>68-70</sup>. However, it has been difficult to generate hypotheses or identify putative regulatory factors just by looking at lists of genes with the highest differential expression level. Housekeeping proteins like RNAP subunits or ribosomal components and obvious metabolic pathways like nucleotide biosynthesis/salvage and phosphate stress dominate these lists. By comparing the transcriptome of *S. venezuelae* treated with RAE non-inducing RNAP inhibitors such as streptolydigin and fidaxomicin alongside rifampin, signatures unique to rifamycins, not just RNAP inhibition, may become more apparent. This approach has the potential to identify proteins or pathways which are involved in induction.

We also made the fortuitous observation that fidaxomicin can antagonize the induction of the RAE (**Figure 5, Supplemental Figure 1**). This compound inhibits RNAP before transcription initiation, meaning that rifampin's steric occlusion of nascent transcripts wouldn't occur on RNAPs inhibited by fidaxomicin and rifamycins. It's tempting to speculate that this could indicate that the production of abortive transcripts (dinucleotides) may be the biochemical cue for induction. We, therefore, wanted to test whether the rifamycin congener, kanglemycin A, which, unlike other rifamycins, blocks the production of these abortive transcripts, is an inducer of the RAE<sup>49,71</sup>. Unfortunately, *Amycolatopsis vancoresmycinia* NRRL B-24208, the producer of kanglemycin A, did not produce this compound in the fermentation media reported in the literature. We tried another 40 different media conditions but could not detect kanglemycin production by high-resolution mass spectrometry or bioactivity assay. Future work should follow up on this line of inquiry, whether by obtaining kanglemycin or by more direct tests such as *in-vitro* transcription in the presence of abortive transcripts or with cell extracts from rifamycin-treated bacteria. The true extent of nucleotide signaling in bacteria is only beginning to be realized, with diverse molecules mediating processes from central metabolism to antiviral immunity<sup>72-74</sup>. As far as I know, the futile synthesis of 2-3nt RNAs is unique to rifamycins, and I would speculate that these products play a role in induction.

## METHODS

### Chemicals and reagents

All chemicals used in this work were purchased from Sigma Aldrich unless stated otherwise. Sorangicin was a gift from Dr. Rolf Müller.

### Bacterial Strains and Conditions

All bacterial species and strains used in this study are listed in Table 1. *S. venezuelae* ATCC 10712 was grown using Bennetts media (10 g potato starch, 2 g casamino acids, 1.8 g yeast extract, and 2 mL of Czapek's mineral mix [10 g KCl, 10 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 12 g NaNO<sub>3</sub>, 0.2 g FeSO<sub>4</sub>, 200 µL concentrated HCl in 100 mL H<sub>2</sub>O] per 1 L H<sub>2</sub>O) prepared with and without 1.5% agar depending on the application. Tryptic soy broth (TSB, BD<sup>®</sup>) was used to grow small cultures of *S. venezuelae* where indicated. Biparental matings were performed on soy flour mannitol agar supplemented with 10 mM MgCl<sub>2</sub>. *Escherichia coli* was grown using LB (Lysogeny broth). *Streptomyces* antibiotic medium (SAM) and R5A media were used for streptolydigin production. SAM contains 15g glucose, 15g soytone, 5g NaCl, 1g yeast extract, 1g CaCO<sub>3</sub>, and 2.5 mL glycerol per litre, pH 6.8. R5A contains 100g sucrose, 0.25g K<sub>2</sub>SO<sub>4</sub>, 10g MgCl<sub>2</sub>·6H<sub>2</sub>O, 10g glucose, 0.1g casamino acids, 5g yeast extract, 21g MOPS, 2mL Trace elements per litre with a final pH of 6.8. Trace elements solution [1mL 1M H<sub>2</sub>SO<sub>4</sub>, 860mg ZnSO<sub>4</sub>·7H<sub>2</sub>O, 223mg MnSO<sub>4</sub>·4H<sub>2</sub>O, 62mg H<sub>3</sub>BO<sub>3</sub>, 125mg CuSO<sub>4</sub>·5H<sub>2</sub>O, 48mg Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O, 48mg CoCl<sub>2</sub>·6H<sub>2</sub>O, 1.8g FeSO<sub>4</sub>·7H<sub>2</sub>O, 83mg KI per litre] . Antibiotics for selection were used at the following concentrations 100 µg/mL ampicillin, 34 µg/mL chloramphenicol, 50 µg/mL

apramycin, 100 µg/mL and 50 µg/mL hygromycin B for *E. coli* and *S. venezuelae* respectively.

### RNA isolation and RT-qPCR

Gene expression data for *rox*, *helR*, and *rgt* were collected simultaneously; however, the data for *rgt* were omitted when we prepared the manuscript, which is reproduced in Chapter 4. The methodology for RT-qPCR for *rgt* is exactly as stated in that chapter, and I will not reproduce them here. The primers used to monitor *rgt* are as 5'GCTACGTACACGCCAGTCAC-3', reverse 5'GTCGATGTCCCACAGCTCC-3'.

**Table 1 – Strains and isolates used in this work**

Strain or Plasmid	Genotype/Use	Source or Reference
<i>Streptomyces venezuelae</i>		ATCC
ATCC 10712		
<i>Escherichia coli</i>		
DH5α	General cloning	Invitrogen
ET12567/pUZ8002:: <i>bla</i>	Methylation deficient strain used for conjugation into <i>Streptomyces</i>	
DH5α pJP3	Microcin J25 production	This work
<b>WAC1561</b>	Streptolydigin production	This work
<b>Plasmids</b>		
pGUS	Apra <sup>r</sup> Spec <sup>r</sup> ; φC31 integrative vector containing promoterless <i>gusA</i>	<sup>34</sup>
pGUS:P <sub>rox</sub>	<i>rox</i> promoter cloned into KpnI/XbaI sites on pGUS	This work
pGUS:P <sub>rox</sub> 18	<i>rox</i> promoter with a 1bp deletion in the spacer region	This work
pGUS:P <sub>rox</sub> 17	<i>rox</i> promoter with a 2bp deletion in the spacer region	This work
pGUS:P <sub>rox</sub> 20	<i>rox</i> promoter with a 1bp insertion in the spacer region	This work
pGUS:P <sub>rox</sub> invert	<i>rox</i> promoter in opposite orientation cloned into KpnI/XbaI sites on pGUS	This work
pJP3	Amp <sup>r</sup> ; Contains microcin J25 biosynthesis operon under IPTG inducible expression.	<sup>46</sup> Gift from Dr. James Link

### Construction of pGUS reporters

The promoter region of *rox* was amplified from *S. venezuelae* genomic DNA using the following primers  $P_{rox}$  FP 5'-ACTGTCTAGACCGCCAGCATGCCCTACCAAC-3' and  $P_{rox}$  RP 5'-TTAGGGTACCGAGAACCGCCCCGTTTCCGC-3' which amplify the intergenic region between *SVEN0480* and *rox*. Underlined bases are XbaI, and KpnI cut sites, respectively.  $P_{rox}$  invert was amplified using the following primers  $P_{rox}$  invert FP 5'-TTAGGGTACCCCGCCAGCATGCCCTACCAAC-3' and  $P_{rox}$  invert RP 5'-ACTGTCTAGAGAGAAACCGCCCCGTTTCCGC-3' PCR products were ligated into the XbaI/KpnI site on pGUS. Spacer mutants were ordered as gBlocks (IDT) with XbaI/KpnI sites and ligated into pGUS. All pGUS constructs were transformed into *E. coli* ET12567/pUZ8002::*bla* for biparental mating into *S. venezuelae*<sup>75</sup>.

### Selection and sequencing of rifampin-resistant *S. venezuelae*

A two-day old culture of *S. venezuelae* pGUS:  $P_{rox}$  was standardized to an OD<sub>600</sub> of 1.0, and 100  $\mu$ L was used to inoculate Bennett's plates containing 40  $\mu$ g/mL of rifampin. *S. venezuelae* containing pGUS vectors sporulate poorly, necessitating mycelial cultures instead of spores. Resistant colonies were re-streaked twice on rifampin-Bennetts plates to ensure purity before preparation of gDNA and mycelial glycerol stocks for storage at -80°C. gDNA was prepared using a PureLink™ Genomic DNA Mini Kit, following the recommended procedure for Gram-positive bacteria. The following primers were used to amplify and sequence the *S. venezuelae* RRDR from gDNA, RRDR<sub>sv</sub> FP 5'-TCGACCACTTCGGCAACCGT-3' and RRDR<sub>sv</sub> RP 5'-TCGATCGGGCACATGCGGCC-3'.

## **GUS assays**

Overnight cultures of *S. venezuelae* grown in Bennetts were diluted to an OD<sub>600nm</sub> = 0.1 and applied as a lawn to agar plates containing 1 in 10 dilutions of Bennetts media supplemented with 80 µg/mL 5-Bromo-4-chloro-3-indoyl-β-D-glucuronide (X-gluc, Alfa Aesar) and appropriate selection. Plates were imaged after 48h incubation at 30°C.

## **Purification of Microcin J25**

*E. coli* DH5α was transformed with pJP3, which encodes for the production of Mcc J25 under the control of an IPTG inducible promoter<sup>46</sup>. An overnight culture of *E. coli* pJP3 was used to inoculate 50mL of LB (a 1/100 dilution) and grown at 37°C with shaking at 250rpm until the OD<sub>600nm</sub> had reached 0.5 at which point production was stimulated by the addition of 1mM IPTG. After three hours of induction, culture supernatants were extracted three times with an equal volume of n-butanol and dried under reduced pressure. 10 mL of water was used to solubilize MccJ25 from the dried material. This mixture was lyophilized, resuspended in 4mL of water, and purified using flash chromatography (CombiFlash, Teledyne) over two 2mL injections on a 13g RediSep C18 column (Teledyne) using a linear gradient of water + 0.1% formic acid (A) and acetonitrile (B). MccJ25 eluted as a single peak between 50-60% B. After lyophilization, we recovered 1.1mg of a fluffy white powder which inhibited the growth of *E. coli* DH5α. We used a high-resolution mass spectrometer (Agilent qTOF 6550 mass detector) to confirm the product as MccJ25. Exact mass calculated for C<sub>101</sub>H<sub>139</sub>N<sub>23</sub>O<sub>27</sub> [M+H]<sup>+</sup> = 2106.0210 Da, observed = 2106.0300 Da (mass error = -4.2ppm).

### **Purification of Streptolydigin**

Spores of WAC1561 were used to inoculate a 50 mL starter culture of SAM. After 5 days of growth, 4L of R5A media was inoculated with 200mL of starter culture (a 1 in 20 dilution) and incubated for 5 days at 30°C with shaking at 250rpm. Cells were removed by centrifugation, and the supernatant and cell pellet were extracted with an equal volume of dichloromethane twice. Organic solvent was removed by rotary evaporation. This crude material was extracted with methanol and streptolydigin was purified on an Agilent 1290 Infinity II preparative HPLC system with an Agilent 5 Prep-C18 column (50x30mm) with the following program. Solvent A = water, solvent B = acetonitrile; 1 minute 5%B, 1-2 minutes 5 to 55%B, 2-14 minutes 55% to 68%B, 14-15 minutes 68% to 95%B, 15-17 minutes isocratic at 95%B, 17-18 minutes 95-5%B, 18-19 minutes 95%B. Flow rate was kept constant at 25mL/min. Streptolydigin eluted from 60-62%B. Following lyophilization, the product was obtained as 20.1 mg of a fluffy off-white powder. Streptolydigin was verified using high-resolution mass spectrometry, exact mass of streptolydigin [M-H] calculated for  $C_{32}H_{44}N_2O_9$  = 599.2968 Da, observed = 599.2950 Da (mass error = -3.0ppm).

### **Susceptibility testing**

Minimum inhibitory concentrations (MIC) were determined by the broth microdilution method as dictated by CLSI protocols with the following modifications<sup>76</sup>. First, *S. venezuelae* was cultured in 3mL of tryptic soy broth (TSB) with 2-3 autoclaved 4mm glass beads to help homogenize the culture. After two days of incubation at 30 °C

and 250rpm, cultures were standardized by OD<sub>600</sub>, and 96 well plates containing *S. venezuelae* were incubated for three days at 30 °C and 250rpm before MIC determination.

### **RNA polymerase purification**

Native RNA polymerase was purified from *S. venezuelae* using the procedure outlined in Kieser *et al.*<sup>75</sup>, which is based on the protocol from Burgess and Jendrisak<sup>77</sup>. A detailed procedure is included in Chapter 4, but the RNAP preparations used here have the following modifications to that protocol. The final Anion exchange steps were excluded, and *in vitro* transcription was performed with pools of all RNAP-containing fractions obtained from the gel filtration step. Pooled and concentrated RNAP was stored at -20°C in 50% glycerol.

### ***In-vitro* transcription**

Multiple round *in-vitro* transcription assays were performed based on the protocol of Kang *et al.*<sup>78</sup>, omitting the addition of heparin which would restrict transcription to a single round. 15µL reactions were prepared by adding template DNA (0.15pmol), RNAP (1.5pmol), and rifampin (various concentrations) to Transcription buffer (40mM Tris pH 7.9, 10mM MgCl<sub>2</sub>, 1.5mM DTT, 0.25mg/mL BSA, 20% (v/v) glycerol). Rifampin was prepared in transcription buffer and added following the addition of template and before the addition of RNA polymerase. Following a 5 minute incubation at 30 °C, transcription was initiated by 3 µL of an NTP mix containing 0.4mM ATP, GTP, UTP, and 0.2mM CTP with 2µCi[α-<sup>32</sup>P] CTP (800 Ci/mmol) (Perkin Elmer) and incubated at 30°C for 2 minutes. Reactions were stopped after 10 minutes by the addition of 50 µL of 375 mM sodium acetate pH5.2, 15mM EDTA, 0.15% SDS, nucleic acids were recovered by precipitation

in ethanol (150 µL) with 10 µg of glycogen added as a carrier. This precipitate was dissolved in loading buffer (7M urea, 0.01 % bromophenol blue in Transcription buffer) and analyzed on a 6% Polyacrylamide gel containing 7M urea. Gels were sandwiched with storage phosphors imaged using a Typhoon®(Amersham).

Template DNA was prepared by PCR using primers to amplify the intergenic region preceding *hrdB* using the primer pairs forward 5'-ATTCGGGCGGCGTGAATGTAC-3' and reverse 5'-CGATCTGCCCATCAGCCTTTCC-3' and for *rox* using forward 5'-CCGTCTCGCGTTCCTCGAAA-3' and reverse 5'-CGTCGGCTCCGTCTCCTTCT-3'. PCR reactions were analyzed for purity on an agarose gel and purified using GeneJET PCR purification kit (Thermo Fisher).

## REFERENCES

1. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
2. O'Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. London: Review on Antimicrobial Resistance, 2016. doi:10.4103/2045-080x.186181
3. Cook, M. A. & Wright, G. D. The past, present, and future of antibiotics. *Sci. Transl. Med.* **14**, (2022).
4. Stokes, H. W. & Gillings, M. R. Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol. Rev.* **35**, 790–819 (2011).
5. Gillings, M. *et al.* The evolution of class 1 integrons and the rise of antibiotic resistance. *J. Bacteriol.* **190**, 5095–5100 (2008).
6. Bennett, P. M. Plasmid encoded antibiotic resistance: Acquisition and transfer of antibiotic resistance genes in bacteria. *Br. J. Pharmacol.* **153**, 347–357 (2008).
7. Waglechner, N., Elizabeth, C. J. & Wright, G. D. Ancient Antibiotics, Ancient Resistance. *EcoSal Plus* **9**, eESP-0027-02020 (2021).
8. Arthur, M., Molinas, C. & Courvalin, P. The VanS-VanR two-component regulatory system controls synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *J. Bacteriol.* **174**, 2582–2591 (1992).
9. Hillen, W. & Berens, C. Mechanisms Underlying Expression of TN10 Encoded Tetracycline Resistance. *Annu. Rev. Microbiol.* **48**, 345–369 (1994).

10. Floss, H. G. & Yu, T.-W. Rifamycin Mode of Action, Resistance, and Biosynthesis. *Chem. rev.* **105**, 621–632 (2005).
11. Campbell, E. A. *et al.* Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **104**, 901–912 (2001).
12. Lin, W. *et al.* Structural Basis of *Mycobacterium tuberculosis* Transcription and Transcription Inhibition. *Mol. Cell* **66**, 169–179.e8 (2017).
13. McClure, W. R. & Cech, C. L. On the mechanism of rifampicin inhibition of RNA synthesis. *J. Biol. Chem.* **253**, 8949–8956 (1978).
14. Murray, J. F., Schraufnagel, D. E. & Hopewell, P. C. Treatment of tuberculosis: A historical perspective. *Ann. Am. Thorac. Soc.* **12**, 1749–1759 (2015).
15. Rothstein, D. M. Rifamycins, alone and in combination. *Cold Spring Harb. Perspect. Med.* **6**, 1–19 (2016).
16. Goldstein, B. P. Resistance to rifampicin: a review. *J. Antibiot. (Tokyo)*. **67**, 625–630 (2014).
17. Surette, M. D., Spanogiannopoulos, P. & Wright, G. D. The Enzymes of the Rifamycin Antibiotic Resistome. *Acc. Chem. Res.* **111**, (2021).
18. Quan, S., Venter, H. & Dabbs, E. R. Ribosylative Inactivation of Rifampin by *Mycobacterium smegmatis* Is a Principal Contributor to Its Low Susceptibility to This Antibiotic. *Antimicrob. Agents Chemother.* **41**, 2456–2460 (1997).
19. Baysarowich, J. *et al.* Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 4886–4891 (2008).
20. Spanogiannopoulos, P., Waglechner, N., Koteva, K. & Wright, G. D. A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7102–7107 (2014).
21. Spanogiannopoulos, P., Thaker, M., Koteva, K., Waglechner, N. & Wright, G. D. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob. Agents Chemother.* **56**, 5061–5069 (2012).
22. Andersen, S. J., Quan, S., Gowan, B. & Dabbs, E. R. Monooxygenase-like sequence of a *Rhodococcus equi* gene conferring increased resistance to rifampin by inactivating this antibiotic. *Antimicrob. Agents Chemother.* **41**, 218–221 (1997).
23. Koteva, K. *et al.* Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem. Biol.* **25**, 403–412.e5 (2018).
24. Spanogiannopoulos, P. Exploring Rifamycin Inactivation from the Soil Microbiome. (McMaster, 2014).
25. Balleza, E. *et al.* Regulation by transcription factors in bacteria: Beyond description. *FEMS Microbiol. Rev.* **33**, 133–151 (2009).
26. Gao, R., Mack, T. R. & Stock, A. M. Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem. Sci.* **32**, 225–234 (2007).
27. Feklistov, A., Sharon, B. D., Darst, S. A. & Gross, C. A. Bacterial sigma factors: A historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* **68**, 357–376 (2014).
28. Kelso, J. Characterizing the mechanism and regulation of a rifamycin monooxygenase in

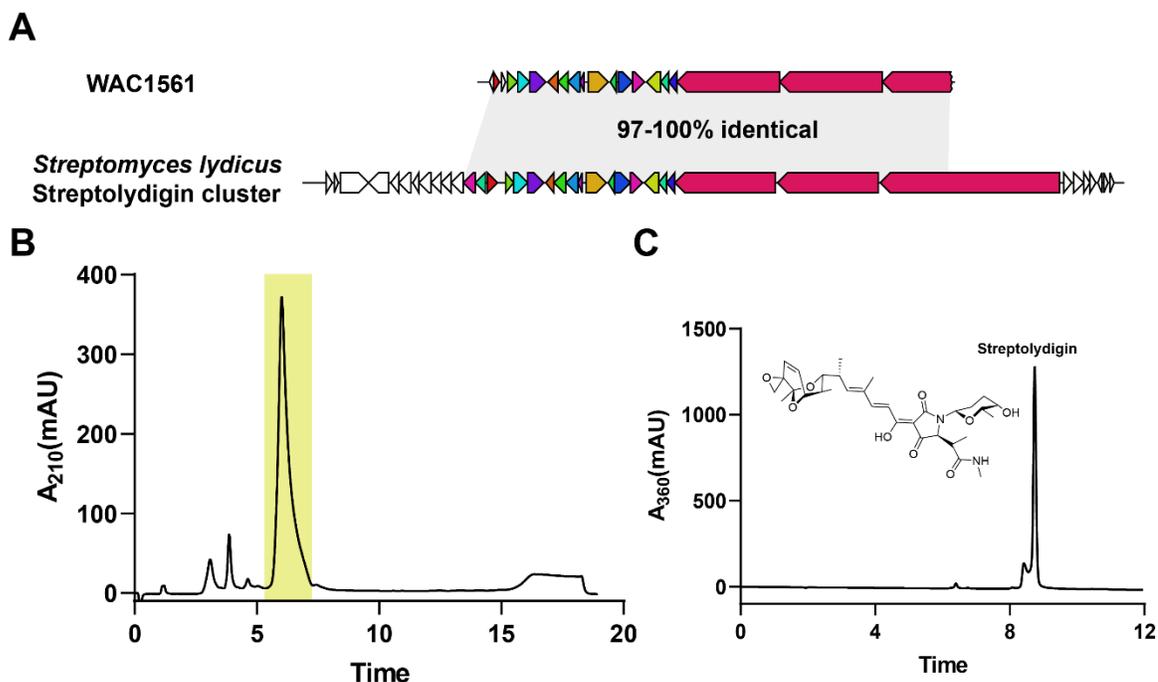
*Streptomyces venezuelae*. (McMaster University, 2016).

29. Moody, M. J., Young, R. A., Jones, S. E. & Elliot, M. A. Comparative analysis of non-coding RNAs in the antibiotic-producing *Streptomyces* bacteria. *BMC Genomics* **14**, (2013).
30. Haugen, S. P., Ross, W. & Gourse, R. L. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6**, 507–519 (2008).
31. Chen, J., Boyaci, H. & Campbell, E. A. Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **19**, 95–109 (2021).
32. Šmídová, K. *et al.* DNA mapping and kinetic modeling of the HrdB regulon in *Streptomyces coelicolor*. *Nucleic Acids Res.* **47**, 621–633 (2019).
33. Strohl, W. R. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **20**, 961–974 (1992).
34. Myronovskyi, M., Welle, E., Fedorenko, V. & Luzhetskyy, A.  $\beta$ -Glucuronidase as a Sensitive and Versatile Reporter in Actinomycetes. *Appl. Environ. Microbiol.* **77**, 5370–5383 (2011).
35. Brown, N. L., Stoyanov, J. V., Kidd, S. P. & Hobman, J. L. The MerR family of transcriptional regulators. *FEMS Microbiol. Rev.* **27**, 145–163 (2003).
36. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber. Lung Dis.* **79**, 3–29 (1998).
37. Jin, D. J. & Gross, C. A. Mapping and sequencing of mutations in the *Escherichia coli* rpoB gene that lead to rifampicin resistance. *J. Mol. Biol.* **202**, 45–58 (1988).
38. Feklistov, A. *et al.* Rifamycins do not function by allosteric modulation of binding of Mg<sup>2+</sup> to the RNA polymerase active center. *Proc. Natl. Acad. Sci.* **105**, 14820–14825 (2008).
39. Campbell, E. A. *et al.* Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674–682 (2005).
40. Boyaci, H. *et al.* Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. *Elife* **7**, 1–19 (2018).
41. Lin, W. *et al.* Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Mol. Cell* **70**, 60–71.e15 (2018).
42. Artismovitch, I., Chu, C., Lynch, S. & Landick, R. A New Class of Bacterial RNA Polymerase Inhibitor Affects Nucleotide Addition. *Science.* **302**, 650–654 (2003).
43. Braffman, N. R. *et al.* Structural mechanism of transcription inhibition by lasso peptides microcin J25 and capistrain. *Proc. Natl. Acad. Sci.* **116**, 1273–1278 (2019).
44. Wilson, K. A. *et al.* Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. *J. Am. Chem. Soc.* **125**, 12475–12483 (2003).
45. Tuske, S. *et al.* Inhibition of bacterial RNA polymerase by streptolydigin: Stabilization of a straight-bridge-helix active-center conformation. *Cell* **122**, 541–552 (2005).
46. Pan, S. J., Cheung, W. L. & Link, A. J. Engineered gene clusters for the production of the antimicrobial peptide microcin J25. *Protein Expr. Purif.* **71**, 200–206 (2010).
47. Pan, S. J., Cheung, W. L., Fung, H. K., Floudas, C. A. & Link, A. J. Computational design of the lasso peptide antibiotic microcin J25. *Protein Eng. Des. Sel.* **24**, 275–282 (2011).

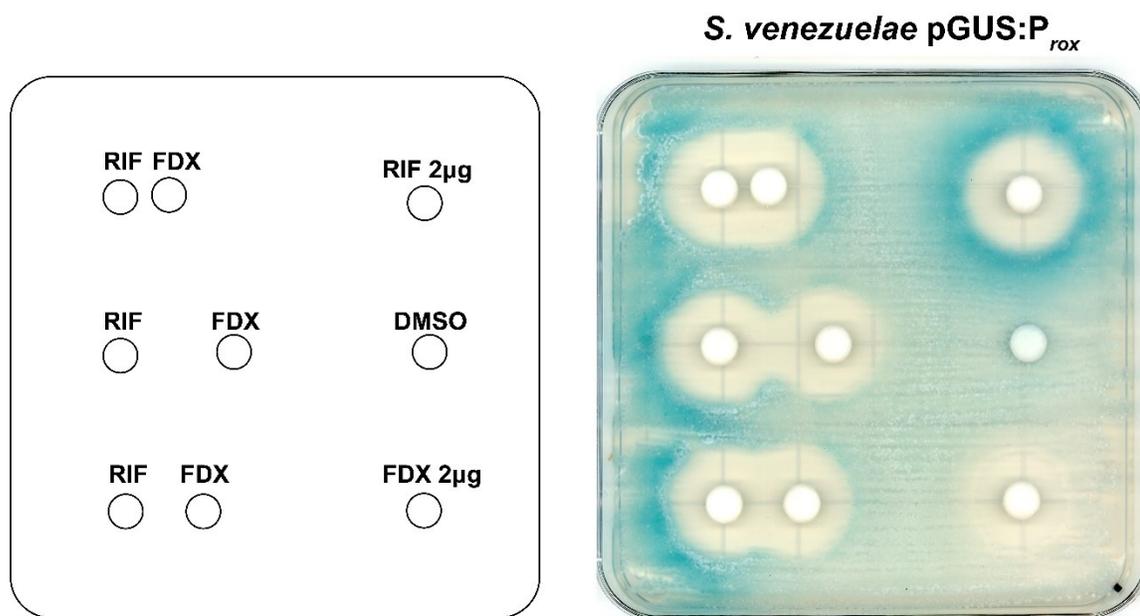
48. Olano, C. *et al.* Deciphering Biosynthesis of the RNA Polymerase Inhibitor Streptolydigin and Generation of Glycosylated Derivatives. *Chem. Biol.* **16**, 1031–1044 (2009).
49. Mosaei, H. *et al.* Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant *Mycobacterium tuberculosis*. *Mol. Cell* **72**, 263–274.e5 (2018).
50. Reznikoff, W. S. The lactose operon-controlling elements: a complex paradigm. *Mol. Microbiol.* **6**, 2419–2422 (1992).
51. Browning, D. F. & Busby, S. J. W. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65 (2004).
52. Sun, F. *et al.* In the *Staphylococcus aureus* two-component system *sae*, the response regulator SaeR Binds to a direct repeat sequence and DNA binding requires phosphorylation by the sensor kinase SaeS. *J. Bacteriol.* **192**, 2111–2127 (2010).
53. Sola-Landa, A., Rodríguez-García, A., Apel, A. K. & Martín, J. F. Target genes and structure of the direct repeats in the DNA-binding sequences of the response regulator PhoP in *Streptomyces coelicolor*. *Nucleic Acids Res.* **36**, 1358–1368 (2008).
54. Hung, D. C. I. *et al.* Characterization of DNA binding sites of the ComE response regulator from *Streptococcus mutans*. *J. Bacteriol.* **193**, 3642–3652 (2011).
55. Körner, H., Sofia, H. J. & Zumft, W. G. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol. Rev.* **27**, 559–592 (2003).
56. Schell, M. A. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**, 597–626 (1993).
57. Stock, A., Robinson, V. & Goudreau, P. Two-Component Signal Transduction. *Annu. Rev. Biochem.* **69**, 183–215 (2000).
58. Lund, P. A. & Brown, N. L. Regulation of transcription in *Escherichia coli* from the *mer* and *merR* promoters in the transposon Tn501. *J. Mol. Biol.* **205**, 343–353 (1989).
59. Parkhill, J. & Brown, N. L. Site-specific insertion and deletion mutants in the *mer* promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. *Nucleic Acids Res.* **18**, 5157–5162 (1990).
60. Martínez-Antonio, A., Janga, S. C. & Thieffry, D. Functional organisation of *Escherichia coli* transcriptional regulatory network. *J. Mol. Biol.* **381**, 238–247 (2008).
61. Becskei Attila & Serrano Luis. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
62. Todor, H. *et al.* Rewiring the specificity of extracytoplasmic function sigma factors. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33496–33506 (2020).
63. Paget, M. & S., M. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules* **5**, 1245–1265 (2015).
64. Koteva, K. *et al.* A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. *Nat. Chem. Biol.* **6**, 327–329 (2010).
65. Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.* **7**, 215–219 (2000).

66. Jacobs, C., Frère, J.-M. & Normark, S. Cytosolic Intermediates for Cell Wall Biosynthesis and Degradation Control Inducible  $\beta$ -Lactam Resistance in Gram-Negative Bacteria. *Cell* **88**, 823–832 (1997).
67. Beckert, B. *et al.* Structural and mechanistic basis for translation inhibition by macrolide and ketolide antibiotics. *Nat. Commun.* **12**, 1–15 (2021).
68. Giddey, A. D. *et al.* A temporal proteome dynamics study reveals the molecular basis of induced phenotypic resistance in *Mycobacterium smegmatis* at sub-lethal rifampicin concentrations. *Sci. Rep.* **7**, 1–13 (2017).
69. Hurst-Hess, K. *et al.* Mycobacterial SigA and SigB cotranscribe essential housekeeping genes during exponential growth. *MBio* **10**, 1–17 (2019).
70. Danelishvili, L. *et al.* *Mycobacterium tuberculosis* Proteome Response to Antituberculosis Compounds Reveals Metabolic “Escape” Pathways That Prolong Bacterial Survival. *Antimicrob. A* **61**, 1–14 (2017).
71. Peek, J. *et al.* Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nat. Commun.* **9**, (2018).
72. Millman, A., Melamed, S., Amitai, G. & Sorek, R. Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. *Nat. Microbiol.* **5**, 1608–1615 (2020).
73. Hengge, R. *et al.* Recent Advances and Current Trends in Nucleotide Second Messenger Signaling in Bacteria. *J. Mol. Biol.* **431**, 908–927 (2019).
74. Irving, S. E., Choudhury, N. R. & Corrigan, R. M. The stringent response and physiological roles of (pp)pGpp in bacteria. *Nat. Rev. Microbiol.* **19**, 256–271 (2021).
75. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. & Hopwood, D. *Practical Streptomyces Genetics*. (John Innes Foundation, 2000).
76. Weinstein, M. P., Patel, J. B., Burnhman, C.-A. & Zimmer, B. L. *CLSI. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically 11th Edition*. (Clinical and Laboratory Standards Institute, 2018).
77. Burgess, R. R. & Jendrisak, J. J. Procedure for the Rapid, Large-Scale Purification of *Escherichia coli* DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography. *Biochemistry* **14**, 4634–4638 (1975).
78. Kang, J. G., Hahn, M. Y., Ishihama, A. & Roe, J. H. Identification of sigma factors for growth phase-related promoter selectivity of RNA polymerases from *Streptomyces coelicolor* A3(2). *Nucleic Acids Res.* **25**, 2566–2573 (1997).

### Supplemental Material



**Supplementary Figure 1 Purification of Streptolydigin.** A) Identification of a streptolydigin biosynthetic gene cluster in WAC1561. Systematic examination of sequenced isolates from the Wright Actinomycete Collection, revealed a 55kb contig from WAC1561 which was almost identical (at the amino acid level) to the characterized cluster from *S. lydicus*. B) Preparative HPLC of WAC1561 culture supernatant extracted with dichloromethane. UV chromatogram showing collected streptolydigin containing fractions in yellow. C) Analytical HPLC demonstrating the purity of streptolydigin.



**Supplementary Figure 2** Fidaxomicin antagonizes induction of the RAE but doesn't display antimicrobial synergy. Discs containing rifampin (RIF) and fidaxomicin (FDX) were placed on top of a lawn of *S. venezuelae* at different distances to observe the effects on killing and induction.

## **CHAPTER 3**

### **A genomic survey of inducible rifamycin resistance**

## **PREFACE**

The work presented in Chapter 3 is a manuscript in preparation for submission.

Surette, M.D., Pawlowski, AC., and Wright, G.D.

MDS designed and performed the experiments, wrote code, performed analysis, made figures, and edited and wrote the manuscript. AP constructed the RAE hidden Markov model. GDW wrote and edited the manuscript.

## ABSTRACT

Rifamycin antibiotics are important medicines with broad-spectrum antibiotic activity, notable for their role in the treatment of *Mycobacterium tuberculosis*. These compounds inhibit bacterial DNA-dependent RNA polymerase, which orchestrates the production of all RNA required for life. Resistance to rifamycins arises relatively rapidly during clinical and laboratory use by selecting mutants with substitutions in the rifamycin binding pocket that diminish affinity for these antibiotics. Outside this context, many environmental Actinobacteria possess a suite of specific and inducible resistance enzymes. Most enzymes work by detoxifying rifamycins directly such as the rifamycin phosphotransferases (Rph), ADP-ribosyltransferases (Arr), glycosyltransferases (Rgt), and monooxygenases (Rox). However, we've also recently characterized an RNAP protection protein, HelR, which is also inducible. Here, we used a conserved DNA motif, the rifamycin-associated element (RAE), required for induction to catalog the inducible rifamycin resistance genes in ~3400 representative Actinobacterial genomes. Although we did not discover any significant new families of resistance genes, this analysis revealed the existing ones to be incredibly abundant. In particular, the actinomycete bacteria are replete with inducible rifamycin resistance mechanisms and the majority are likely capable of rifamycin inactivation. We also map the rifamycin resistance mechanisms in the medically important *Mycobacterium* genus, which can potentially inform drug development efforts. Lastly, we noted that many bacteria harbor multiple RAEs, and deduced that these are most often present in HelR-inactivating enzyme combinations, highlighting the critical role of HelRs in rifamycin resistance. This analysis offers a unique and comprehensive view of antibiotic resistance at the phylum level.

## INTRODUCTION

Rising rates of antibiotic resistance have jeopardized some of our most critical medicines. Previously treatable pathogens have acquired resistance to many frontline agents, and the deaths from these infections are already mounting<sup>1</sup>. In 2019 alone, antibiotic resistance was responsible for 1.4 million deaths worldwide and implicated in another 4 million<sup>2</sup>. Soil microbes are the original inventors of most classes of clinically used antibiotics. In this context, antibiotics pre-date anthropogenic use by millennia, offering environmental microbes abundant opportunities to evolve strategies to overcome these compounds<sup>3,4</sup>. Moreover, this reservoir of resistance can compromise the efficacy of clinical antibiotics through the lateral transfer of resistance genes from non-pathogenic species to pathogens. Identifying and understanding resistance mechanisms in all bacteria, regardless of their immediate clinical relevance, is crucial because their existence in the environment poses a potential threat to their clinical efficacy.

Rifamycins are broad-spectrum antimicrobial agents and World Health Organization essential medicines<sup>5</sup>. Clinically, these antibiotics are used mainly in the treatment of *Mycobacterium tuberculosis*, and in this context, rifamycin resistance is well understood at the molecular level. Rifamycins inhibit the growth of bacteria through a tight interaction with prokaryotic DNA-dependent RNA polymerase (RNAP) that blocks the synthesis of full-length transcripts, thereby preventing growth<sup>6-8</sup>. Spontaneous mutations give rise to substitutions in the rifamycin binding pocket on the  $\beta$ -subunit of RNAP, which prevent rifamycin binding and confer high-level resistance<sup>9</sup>. The relative frequency of these mutations is the primary barrier to rifamycin use in other medicinal contexts.

Rifamycins are natural products, initially isolated from *Amycolatopsis mediterranei*, but are now known to be produced by bacteria from diverse families within the Actinobacteria. The natural origin of these antibiotics has unsurprisingly resulted in many rifamycin-resistant soil organisms. What is surprising is the *mechanisms* of resistance these bacteria employ. Instead of relying on substitutions in RpoB, Actinobacteria are known to produce four distinct families of rifamycin-inactivating enzymes, phosphotransferases (Rph), glycosyltransferases (Rgt), ADP-ribosyltransferases (Arr), and the hydroxylases (Rox/Iri)<sup>10</sup>. These enzymes transfer chemical diverse groups to several distinct moieties on the rifamycin scaffold, preventing them from binding RNAP. More recently, we characterized a novel mechanism of resistance where the RNAP protection protein, HelR, displaces rifamycins from the enzyme (Chapter 4)<sup>11</sup>. Remarkably, in Actinobacteria, expression of all 5 enzymes is induced by rifamycin antibiotics. The mechanism of this regulation is not well understood, but induction depends on a highly conserved *cis*-regulatory DNA sequence, the rifamycin-associated element (RAE)<sup>12</sup>. The RAE was first characterized as a 19bp inverted repeat (two 9bp repeats with a 1bp spacer). However, in the previous chapter, we identified a highly conserved -10 site spaced precisely 10bp downstream from the 3' end of the RAE and 19bp away from a putative -35 site within the RAE. We termed the larger motif, which encompasses these additional sequences, the extended RAE. In several instances, we have used the RAE sequence to guide our search for novel mechanisms of rifamycin resistance, such as Rph and HelR. However, it has been challenging to comprehensively characterize genes regulated by the RAE due to its directionless nature. The extended RAE, on the other hand, points towards

the ORFs it controls, enabling us to survey microbial genomes for this motif systematically and to identify the protein(s) it regulates.

In this study, we built a hidden Markov model (HMM) for the extended RAE to identify all the RAE-associated genes in almost 3400 representative Actinobacterial genomes. We used this dataset to provide the first estimate for the prevalence of inducible rifamycin resistance, identify the families of proteins regulated by the RAE, and investigate the taxonomic distribution of specific resistance mechanisms among Actinobacteria, including pathogenic mycobacteria. Lastly, we also examined the frequency of genomes that encode multiple RAEs. We demonstrated that these preferentially encode combinations of a single HelR and >1 interchangeable inactivating enzymes, highlighting their different yet complementary roles in resistance.

## Results

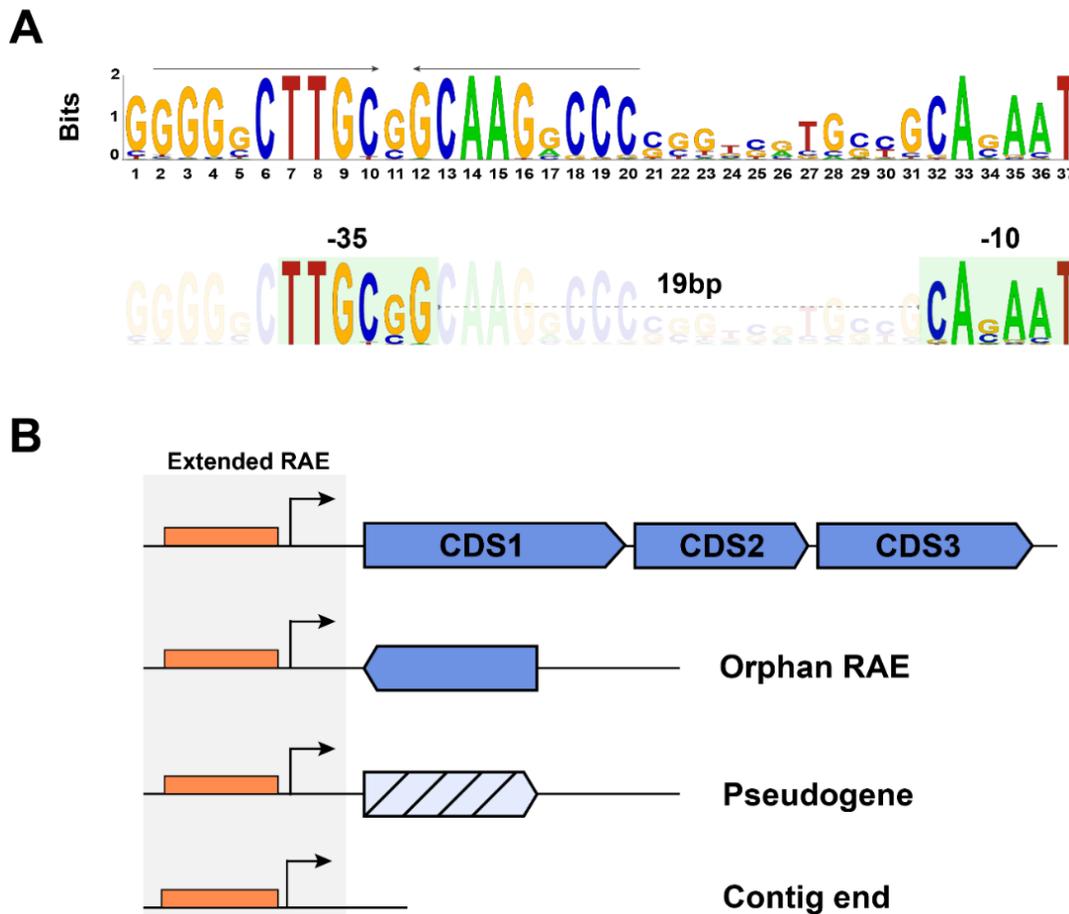
We began by validating the extended RAE and building a HMM that we could use to search nucleotide databases. Using BLASTn, we identified RAE sequences from Actinobacteria and pulled the DNA 100bp upstream and downstream from each RAE. We used an alignment of these sequences to build a 37bp model of the extended RAE (**Figure 1A**). The palindromic sequence predicted here is 21bp long instead of 19bp, including a run of 4 consecutive guanines at the 5' end with relaxed symmetry at the 3' end CCC(G/C). At the 3' end of the model lies the predicted -10 site (CA(G/C)AAT). Promoter searches have traditionally been challenging bioinformatically<sup>13,14</sup>; the variable length of the spacer motif is a specific problem when building HMMs, as these models do not handle gaps or

inserts well. The strictly conserved spacing of the extended RAE makes it uniquely well-suited to the HMM approach.

Using this model, we wanted to query Actinobacterial genomes to catalogue the genes controlled by the RAE and determine their relative abundance and distribution in Actinobacteria. Public repositories such as Genbank or RefSeq have, in some instances, thousands of genomes from single species such as *M. tuberculosis* (>7000) or *Mycobacterium abscessus* (*Mycobacteroides abscessus*, >1800), which makes relative abundance challenging to interpret. We used the NCBI collection of representative genomes from the phylum Actinobacteria (Actinobacteriota) to avoid this problem<sup>15</sup>. This collection contains the best quality genome for each species with standing in the literature, offering a representative overview of diversity within the phyla. While there is still room for bias within this dataset, it does prioritize actual observed phenotypic, morphological, and genetic diversity while limiting itself to just one genome per species.

We queried representative Actinobacterial genomes using our model for the RAE and categorized the downstream sequences as follows. Immediately downstream coding sequences (CDSs) were designated CDS1, and genes in the CDS2 and 3 positions were also analyzed, provided they were in the correct orientation. We became aware of *rgt-helR* operons which seemingly contain two genes controlled by a single RAE, and didn't want to miss other operons. We called extended RAEs upstream of genes on the wrong strand orphan RAEs, and we used these as a rough approximation of our false discovery rate. Some orphan RAEs are likely genuine, but a preponderance of these would imply that our

model is too relaxed. Pseudogenes were placed in a separate category, as were RAEs found upstream of a contig end (with no CDS). This query returned 3328 RAEs from 3374 genomes, almost precisely 1 per genome on average. We identified 3198 CDS1s, 2215 sequences at the CDS2 and 3 positions, 100 pseudogenes, 23 orphan RAEs, and 7 found near contig ends. Orphan RAEs corresponded to just 0.69% of our total, indicative of stringent search criteria. Our previous experience isolating rifamycin-inactivating bacteria suggested this trait is relatively rare, but the RAE is remarkably abundant in Actinobacteria.

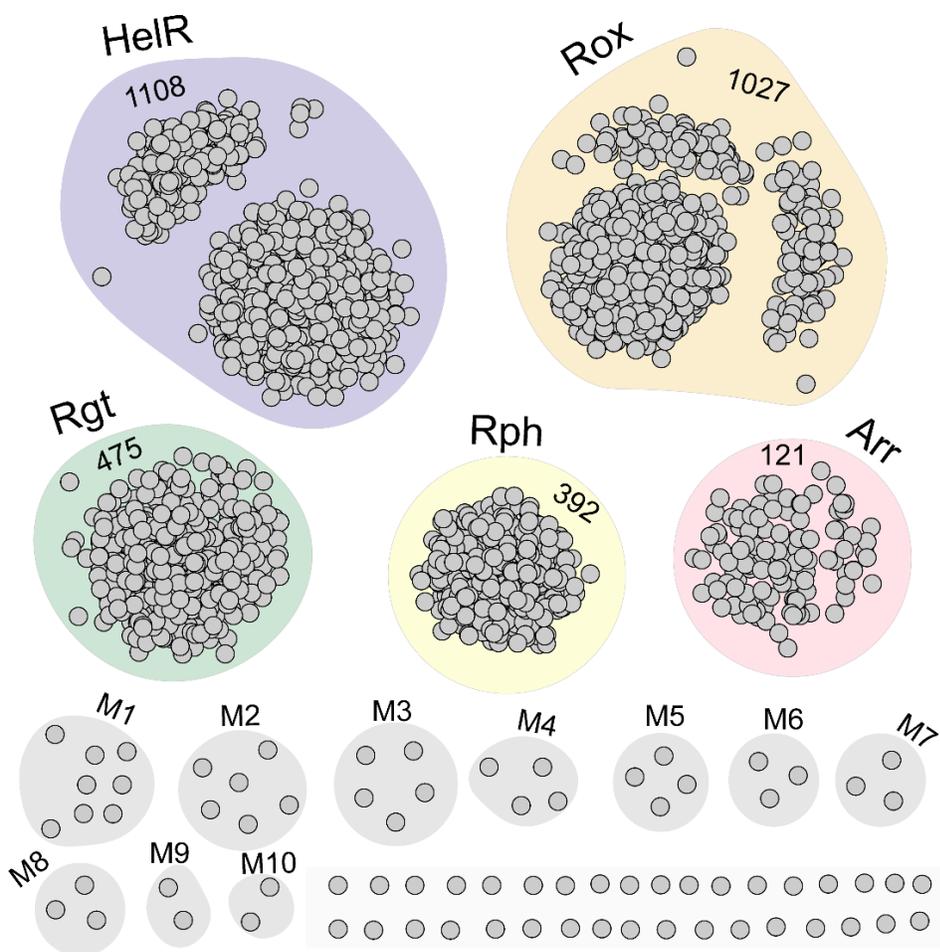


**Figure 1. The extended RAE. A)** Consensus sequence of the extended RAE is represented as a sequence logo. Arrows represent the initial inverted repeats; the putative promoter elements are shown below. **B) Classification scheme for RAE-associated genes and features**

To categorize RAE-associated genes (RAGs), we performed an all against all BLAST, sorting these proteins into broad families based on sequence homology. We depicted the results for CDS1 as a sequence similarity network (SSN) in **Figure 2**, clusters corresponding to known rifamycin resistance genes were named accordingly. The helicase-like protein associated with the RAE (HelR) is the most common RAE-associated gene in Actinobacteria, followed closely by Rox. These two proteins make up the majority of RAE-associated genes. In descending order of abundance, the remaining known resistance enzymes are Rgt, Rph, and Arr. Apart from these known protein families, we identified 10 small clusters with between 8 and 2 representatives; we termed these miscellaneous genes 1 through 10 (**Figure 2**). Lastly, 34 proteins were the sole member of their cluster. Of the 3197 CDS1 proteins, only 74 (2.2%) clustered outside the inactivating enzymes or HelR, suggesting that the major families of rifamycin resistance enzymes have already been characterized.

Nevertheless, the miscellaneous proteins could be genuine but not widespread rifamycin resistance genes, perhaps being recent products of regulatory capture by the RAE. We queried the NCBI's conserved domain database using a member of each cluster to evaluate the potential function of these proteins (**Table 1**)<sup>16</sup>. Some have the clear potential to confer rifamycin resistance, such as M1 (n = 8), which encodes a major facilitator family (MFS) efflux pump. Curiously, M1 is a specific match to a yeast efflux pump. M7 encodes flavin utilizing monooxygenases and could potentially perform a Rox-like reaction. M3 and M6 encode P420 (an atypical flavin cofactor) utilizing enzymes. These catalyze various reactions, including hydroxylations, oxidations, and reduction.

Members of this family have well-characterized roles in vitamin B<sub>6</sub> and oxytetracycline biosynthesis, which both possess aromatic and/or unsaturated ring systems similar to rifamycins indicating it could be a potential substrate<sup>17,18</sup>. The remaining genes don't have apparent ties to rifamycins or known resistance mechanisms. Future work will be required to determine whether any representatives from these less abundant protein clusters can confer resistance.



**Figure 2 Sequence similarity network of RAE-associated protein families.** All 3198 RAE-associated proteins are represented as a network of nodes (unique proteins) grouped by sequence similarity at the amino acid level. Families composed of characterized rifamycin resistance genes are labelled accordingly. Miscellaneous clusters, which cluster outside the known genes are denoted M1-M10.

**Table 1 – Miscellaneous RAE-associated genes**

<b>Family</b>	<b>n</b>	<b>CDD*</b>	<b>Name</b>	<b>E-value</b>	<b>Function</b>
<b>M1</b>	8	Cd17502	MFS_Azr1_ MDR_like	7.36e <sup>-85</sup>	Efflux Pump
<b>M2</b>	6	COG0596	MhpC	4.65e <sup>-27</sup>	Pimeloyl-ACP methyl ester carboxylesterase
<b>M3</b>	5	Cl25688	Rv1155_F420 superfamily	4.87e <sup>-17</sup>	F420-dependent enzyme
<b>M4</b>	4	COG0262	FolA	1.87e <sup>-22</sup>	Dihydrofolate reductase
<b>M5</b>	4	PRK03592	PRK03592	0	Haloalkane dehalogenase
<b>M6</b>	3	TIGR04023	PPOX_MSMEG _5819	2.99e <sup>-59</sup>	Hydroxylase (FMN/PLP)
<b>M7</b>	3	Cl19096	Flavin utilizing monooxygenases Superfamily	2.24e <sup>-55</sup>	FMN Monooxygenase
<b>M8</b>	3	COG0488	Uup	0	ATPase of ABC transporters with duplicated ATPase domains
<b>M9</b>	2	Cl27690	Mycofact_TetR superfamily	4.22e <sup>-22</sup>	TetR – Mycofactocin biosynthesis
<b>M10</b>	2	None	N/a	N/a	N/a

\*Top hit from the conserved domain database (NCBI)

In contrast to the results for CDS1, CDS2 and 3 were incredibly diverse, harboring dozens of small protein clusters ( $n < 20$ ) and hundreds of single-member clusters (**Supplementary Figure 1, Supplementary Table 1**). This most likely indicates that multigene operons controlled by the RAE are not widespread, and this approach is flagging

proteins unrelated to the RAE. For instance, the largest cluster encodes HelRs, which we know from the presence of *rgt-helR* operons in various *Streptomyces* spp., are genuine. This cluster (n = 177) contains an order of magnitude more proteins than the next largest cluster (Rox enzymes, n = 18). **Supplementary Table 1** contains the conserved domains identified in the next 15 most abundant families (16-9 members). Apart from a family of FAD-dependent oxidoreductases (potential Rox-like enzymes, M4-2) and another family of MFS efflux pumps (M12-2), these proteins possess no obvious relationship to rifamycin resistance. As with the miscellaneous protein families identified in CDS1, some of these genes may represent novel resistance mechanisms; however, they are not abundant and, therefore, not significant elements of the rifamycin resistome.

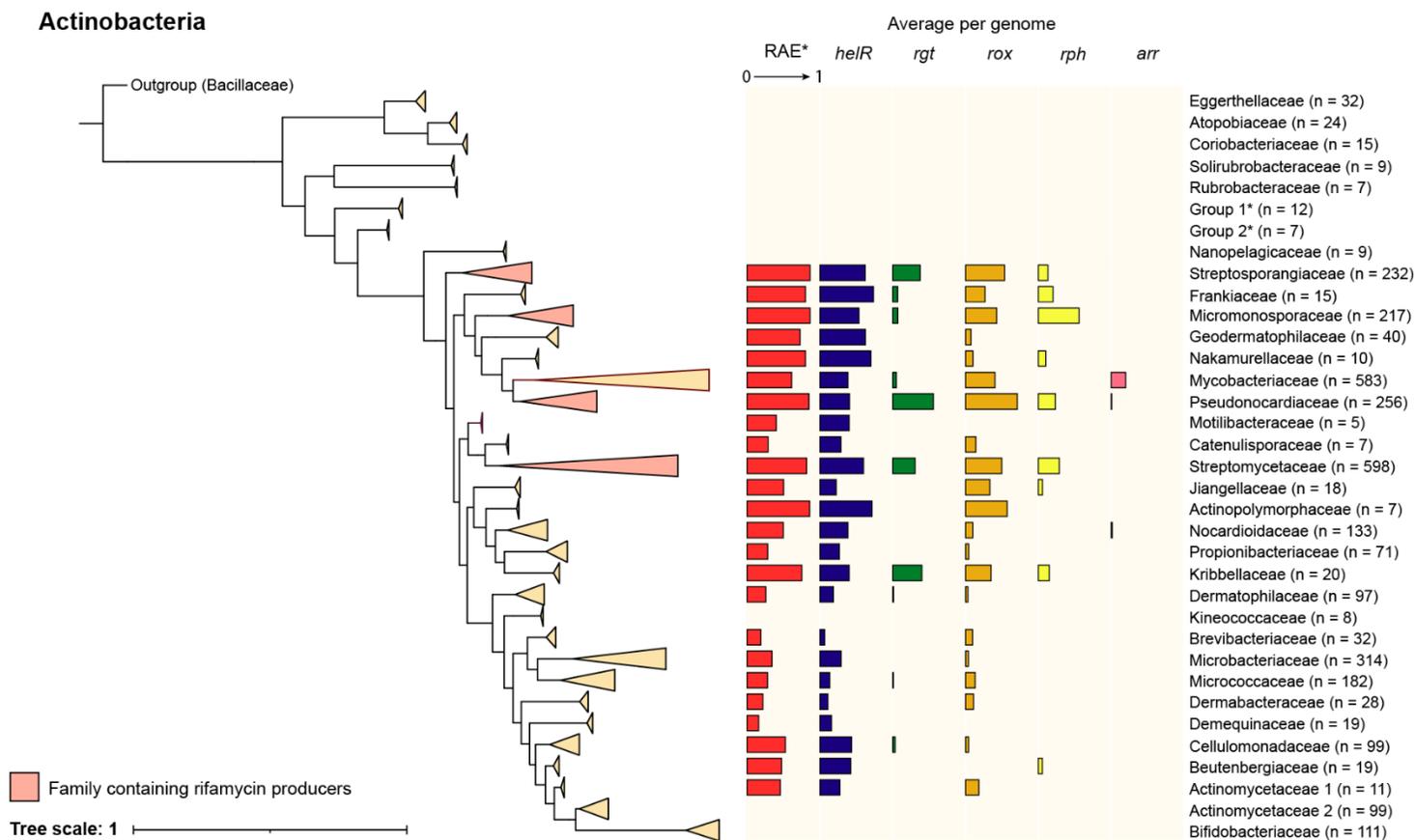
Induction of the RAE loci by rifamycin antibiotics is still enigmatic. We initially anticipated that this search could reveal regulatory proteins essential for induction. Two proteins in the CDS1 position, corresponding to cluster M2, are predicted to be TetR family regulators (**Table 1**). Two more families were also identified in the CDS2,3 positions (**Supplementary Table 1**), encoding YjvK and CitB family regulators, respectively. We determined using BLAST that these proteins are not conserved across all Actinobacteria known to have functional RAEs, and are, therefore, unlikely to play any role in induction.

To assess the phylogenetic distribution of RAEs and associated genes, we took our 3375 Actinobacterial genomes and used the genome-based taxonomy database toolkit (GTDBtk) to obtain detailed taxonomic classifications for each genome and construct a phylogenetic tree<sup>19,20</sup>. NCBI's taxonomy relies on user-supplied information, which can lead to misclassification. For instance, GTDBtk did identify a single Firmicute genome

within our set of representative Actinobacteria. Using the protein clusters identified in **Figure 2**, we assigned each RAE-associated CDS to a protein family (Ie. HelR, Rox, Rgt, Rph, Arr, Miscellaneous) or alternative (pseudogene, orphan, or contig end). We omitted CDS from the 2 and 3 positions for this analysis except for HelR and Rox. Using these assignments, we could quantify the RAE-associated ‘inventory’ of each genome. For example, *Mycobacterium smegmatis* MC<sup>2</sup>155 encodes 1 HelR, 1 Rox, 1 Arr, 0 Rph, 0 Rgt, 0 miscellaneous, 0 pseudogenes, 0 orphans, and 0 contig end associated RAEs. We compiled this data at the family level and displayed the proportion of family members that possessed at least one RAE and the proportion of family members encoding a RAE associated copy of HelR, Rgt, Rox, Rph, Arr on a phylogeny containing all families with at least 5 members (**Figure 3**). This family-level data revealed several salient features. Firstly, the RAE is absent from anaerobic Actinobacteria such as Bifidobacteriaceae, Eggerthellaceae, and Actinomycetaceae 2 (*Actinomyces* and related genera, **Supplementary Table 2**). GTDB places *Georgenia*, *Bogoriella*, and *Oceanitalea* spp. into Actinomycetaceae, and they form a separate branch on the phylogenetic tree in **Figure 3**, which I’ve designated Actinomycetaceae 1<sup>20</sup>. Unlike most Actinomycetaceae 2, these genera are capable of aerobic growth and possess all the RAEs in the family. The RAE is also absent from deeply branching families such as the Rubrobacteraceae, Solirubrobacteraceae, Atopobiaceae, Ca. Nanopelagicaceae, and the Coriobacteriaceae. Rifamycin producers are known to be saprophytic, mesophilic aerobes. Accordingly, it is unlikely that anaerobic organisms would have encountered rifamycins until human use began in the 1960s. Likewise, the deeply branching families include thermophiles

(Rubrobacteraceae)<sup>21</sup>, marine oligotrophs (Ca. Nanopelagicaceae)<sup>22</sup>, and facultative anaerobic members of the human microbiome (Atopobiaceae, Coriobacteriaceae)<sup>23,24</sup>. The RAE is most abundant in the so-called actinomycetes and related genera. It appears to be particularly common in families with known rifamycin producers (in this case, rifamycins include the related compounds streptovaricin and chaxamycin). We detected no rifamycin producers that contained RAEs, most likely because their insensitive RNAP provides sufficient resistance, and inactivating their own antibiotics would be counter-productive. For example, within the genus *Amycolatopsis* (n = 71), only 15 representatives lacked a RAE, and 10 were rifamycin producers (**Supplementary Figure 2**). The data presented here is superficially consistent with the notion that inducible rifamycin resistance is a trait that evolved in bacteria exposed to these compounds in their natural environment.

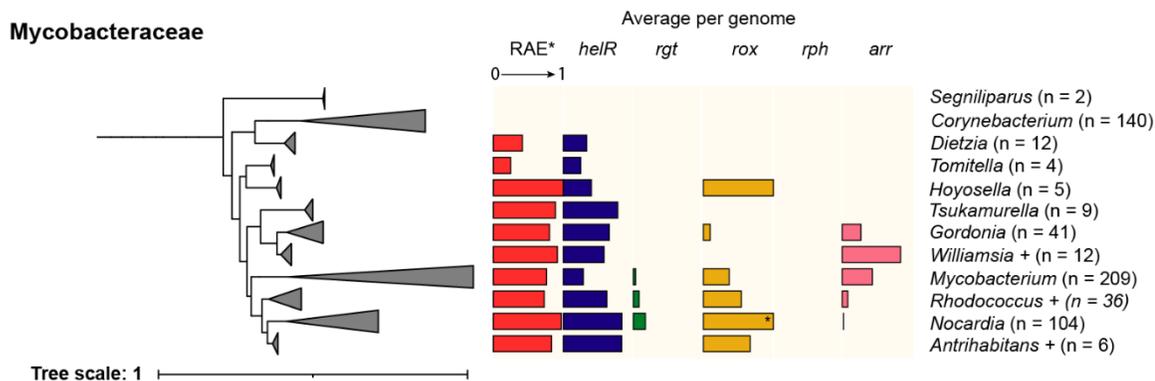
The relative abundance of specific rifamycin resistance genes also displays clear taxonomic trends. HelR is the most abundant RAE-associated protein in our dataset and the most widespread phylogenetically. It is present across all families in **Figure 3** that harbor RAEs. Rox enzymes, which are nearly as abundant as the RAE, share a similarly broad distribution. They are particularly abundant in the rifamycin-producing families, particularly Pseudonocardiaceae. The other resistance genes show clearer taxonomic distributions. For instance, Rgt is an abundant member of just four families. Rph is present at a rate of over 0.5 per genome within Micromonosporaceae, double the abundance in any other family. Arr, the rarest of the inactivating enzymes, is closely associated with the Mycobacteraceae, 116 of 121 come from this family.



**Figure 3 Distribution of the RAE and associated resistance mechanisms in Actinobacteria.** The average number of *HelR*, *Rgt*, *Rox*, *Rph*, *Arr* per genome for every family with greater than 5 members is shown as a horizontal bar graph corresponding to that family's location on a phylogenetic tree of the Actinobacteria. Average per genome ranges from 0 to 1. \*RAE denotes the average number of family members with  $\geq 1$  RAE. Groups 1 and 2 contain monophyletic groups of smaller families. Group 1 is the Acidimicrobiaceae, Ilumatobacteraceae, Microtrichaceae, SKKL01, and UBA8139. Group 2 includes the Nitriliruptoraceae, Egibacteraceae, and Euzebyaceae. For a list of genera in Actinomycetaceae 1 and 2 (see **Supplementary Table 2**). Tree was visualized using iTOL<sup>25</sup>.

Overall, the wide phylogenetic distribution of the RAE and associated resistance genes implies an ancient origin, extensive horizontal transfer, or both. Regardless of the evolutionary history, inducible rifamycin resistance appears to be a widespread and relatively common trait in many Actinobacteria.

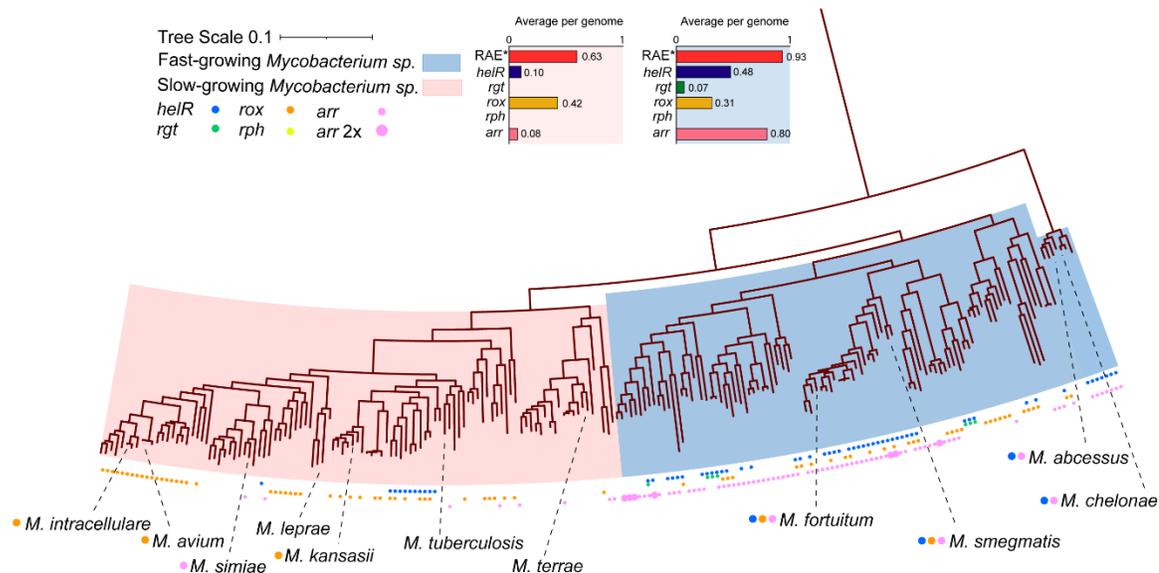
We decided to examine the inventory of RAE-associated genes in mycobacteria in more detail, as rifamycins are frontline agents against several important *Mycobacterium* spp. Within the family Mycobacteraceae, we observed considerable variability in RAE frequency and resistance enzyme abundance (**Figure 4**). All 140 members of the *Corynebacterium* lack a RAE, suggesting that inducible rifamycin resistance has probably been lost since it split from the rest of the Mycobacteraceae. Inducible rifamycin resistance is present in all but one other genus within the family and adjacent families (**Figures 3 and 4**). On average, 97% of *Nocardia*, which includes several opportunistic pathogens, encode at least one RAE. They average one of the highest ratios of HelR for any populous genera at >0.8 per genome while also encoding a significant number of Rgts and a staggering 1.2 Rox enzymes per genome, indicating widespread redundant copies of this enzyme. This extensive armament of rifamycin-resistance genes is curious because many *Nocardia* species encode a second copy of *rpoB* which carries a rifamycin-resistant allele and is known to confer high-level resistance<sup>26</sup>. The selective pressure required to acquire and maintain many resistance mechanisms must be considerable. The Mycobacteraceae lack Rph enzymes entirely and encode very few Rgt enzymes. Conversely, this family contains all but 5 Arrs detected by our analysis, and even within this family, they are not widely distributed. They primarily belong to *Mycobacterium* and *Gordonia* spp. In comparison,



**Figure 4 RAE and rifamycin resistance gene frequency in the Mycobacteraceae.** *Williamsia+* contains *Williamsia A*, *Gordonia B*, and *Williamsia*. *Rhodococcus +* contains *Rhodococcus*, *Rhodococcus B*, and *Rhodococcus C*. GTDB splits into many common genera, often terming them A, B, C, etc.

*Williamsia+* possess a high proportion of Arrs on average but is not a numerous group (n = 12). In total, 90 of 121 RAE-associated Arrs fall within the genus *Mycobacterium*, the narrowest taxonomic distribution among the mechanisms associated with the RAE. This distribution suggests that Arr was recently appropriated for rifamycin resistance within this family.

Within the medically important genus *Mycobacterium*, inducible rifamycin resistance is abundant. Note that the genome-based taxonomic approach used here suggests a re-unification of the *Mycobacterium* genus; we'll be referring to the taxonomic classifications put forth by GTDBtk. Still, readers should be aware that this is an area of ongoing debate in the literature<sup>27,28</sup>. Mycobacteria fall into two distinct groups based on their growth kinetics, the fast growers and the slow growers (**Figure 5**). The major human pathogen, *M. tuberculosis*, belongs to the slow-growing clade, although mycobacteria from both can cause disease. The abundance of RAEs and RAGs is significantly higher in the



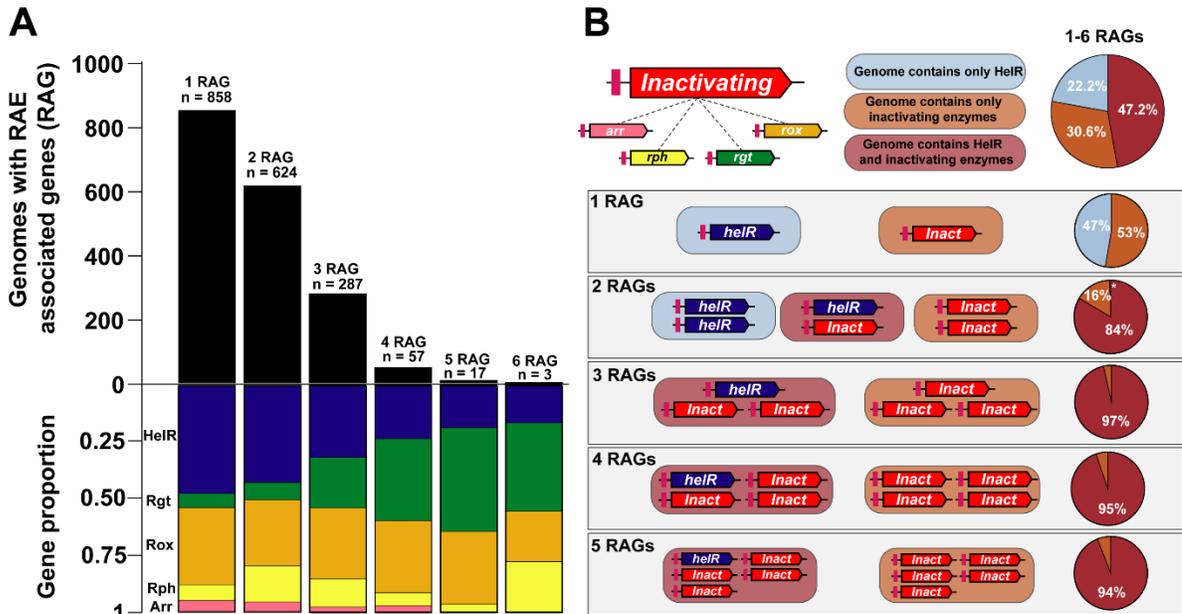
**Figure 5 Inducible rifamycin resistance in the mycobacteria.** A phylogeny of the *Mycobacterium* spp. with any RAE-associated resistance genes indicated at the tips. Many disease-causing species harbor inducible resistance mechanisms. A graph summarizing the average RAEs and RAGs per genome broken down by fast vs. slow-growing species is above the phylogeny. \*RAE average is computed for genomes that encode  $\geq 1$  RAE.

fast-growing species; 93% of fast growers encode at least one RAE as opposed to just 63% of slow growers, with a far higher frequency of multiple RAEs observed in the fast growers. They also differ in their RAE-associated gene cargo; the fast growers possess, on average, 0.8 Arr's per genome compared to 0.08 in the slow growers (**Figure 5**). The Rox enzyme, on the other hand, was more abundant in the slow growers (0.42 vs. 0.31). *Mycobacterium abscessus* and *M. smegmatis* are well-known examples of fast-growing species with intrinsic rifamycin resistance<sup>38</sup>. Here, we demonstrate that rifamycin-resistance genes are widespread within the genera.

RAE-associated rifamycin resistance is a widespread trait in Actinobacteria, and this seems to be true even for disease-causing *Mycobacterium* spp. (**Figure 5**). The obligate

pathogens *M. tuberculosis* and *Mycobacterium leprae* both lack RAEs entirely, but apart from *Mycobacterium terrae*, all other species harbored RAE-associated rifamycin resistance mechanisms. The opportunistic pathogens, *Mycobacterium intracellulare*, *Mycobacterium avium*, and *Mycobacterium kansasii*, all encode Rox enzymes. Fast-growing opportunistic pathogens *M. abscessus*, *Mycobacterium chelonae*, and *Mycobacterium fortuitum* encode HelR/Arr or HelR/Arr/Rox combinations.

During this work, we frequently observed bacteria encoding 2-3 RAEs and wanted to investigate this phenomenon in more detail. We quantified the number of RAE-associated genes (RAGs) per individual bacterial genome ranging from 0 to 6 (**Figure 6A**). For this analysis, we omitted miscellaneous genes, pseudogenes, and RAEs with no associated CDS (excluding ~6% of the RAEs we identified). For organisms with at least one RAG, multiple RAGs are more common than singular ones (median = 2 RAGs). Moreover, filamentous saprophytes and Mycobacteriaceae encode 2+ RAGs more frequently than other families (**Supplementary Figure 3**). We noticed that as an organism encodes more RAGs, the proportion of *helRs* declines (**Figure 6A**). We rationalized this with the hypothesis that organisms encode only a single *helR*, and all other RAGs are devoted to inactivating enzymes. We binned all four inactivating enzymes together and analyzed the frequency of genomes encoding 1) *helR* alone, 2) inactivating enzymes alone, and 3) *helR*-inactivating gene combinations for all genomes with a RAE (n = 1846) (**Figure 6B**). Across all genomes with a RAG, *helR*-inactivating gene combinations were the most common, making up most combinations found with 2+ Rags. A single RAG must be insufficient for many bacteria to confer the level of resistance required in their

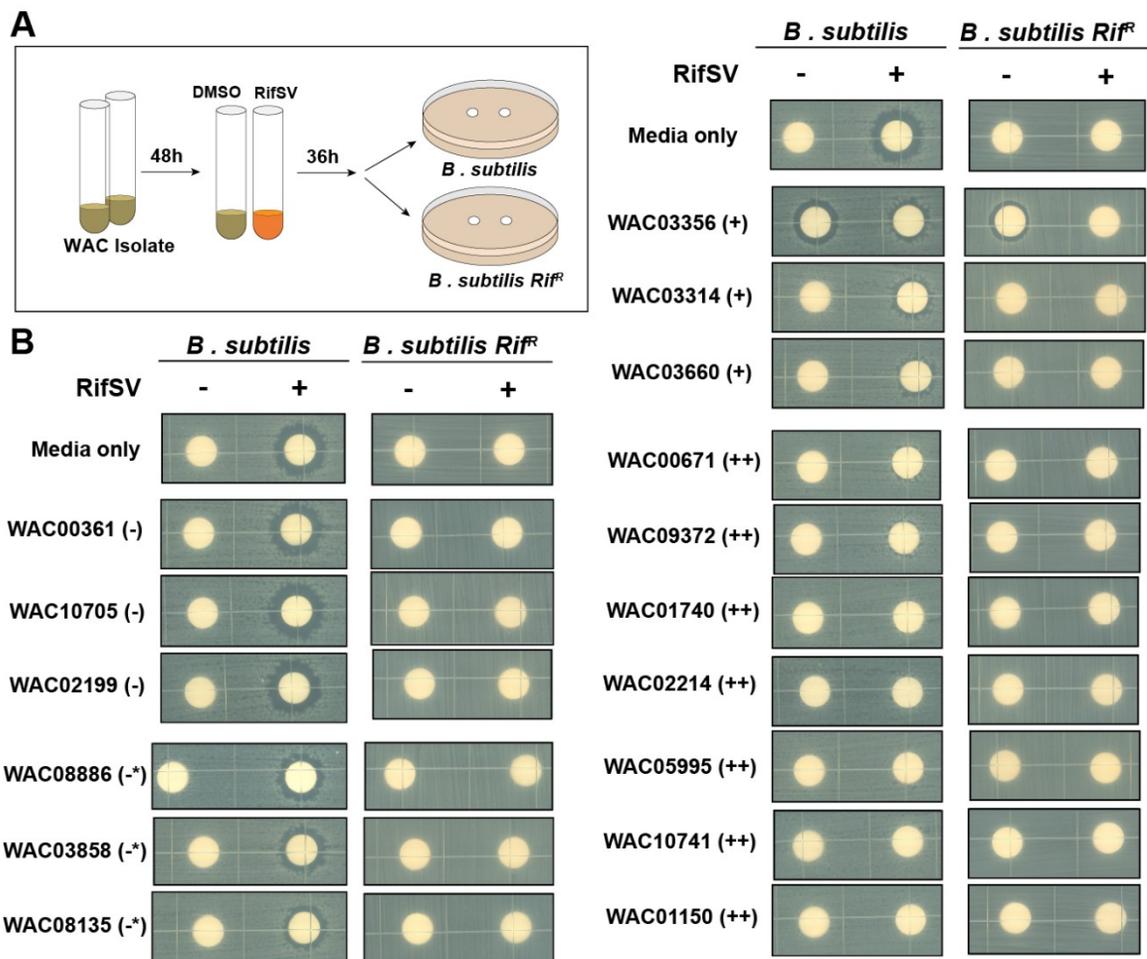


**Figure 6 A) Prevalence and specific inventory of multi-RAE organisms.** (Top) Number of genomes with 1-6 RAE associated genes (RAGs), in this case HelR, Rgt, Rox, Rph, Arr (not miscellaneous genes). (Bottom) The proportion of each RAG from organisms with 1-6 RAGs. **B) Actinobacteria favour combinations of HelR and inactivating genes.** The proportion of genomes that contain HelR only, inactivating enzymes only, and a combination of inactivating genes and a HelR are shown for all organisms with a RAG (top) and 1-5 RAGs individually. Genomes with 6 RAGs had 100% HelR and inactivating enzyme combinations (not shown).

environment. This redundancy is not random and shows a strong bias towards arming bacteria with combinations of HelR and inactivating enzymes. Based on the mechanism of HelR discussed in Chapter 4, this affords bacteria an enzyme to dissociate rifamycin-RNAP complexes coupled with inactivating enzymes that detoxify the antibiotic and prevent re-binding.

Our genomic survey has revealed that rifamycin-inactivating genes associated with the RAE are ubiquitous in common genera such as *Streptomyces*. Previously we've assayed almost 500 actinomycetes for rifampin resistance and inactivation. Using a resistance

cutoff of 20  $\mu\text{g}/\text{mL}$ , approximately 10% of the collection was resistant, and about half of the resistant strains could inactivate rifampin<sup>29</sup>. Our measurement of rifamycin inactivation in ~5% of actinomycetes stands in stark contrast to the 70-80% of Streptomycetaceae, Streptosporangaceae, Micromonosporaceae, and Pseudonocardaceae predicted by this work (**Supplementary Figure 4**). We hypothesized that the resistance cutoff eliminated the vast majority of rifamycin inactivators in the previous study. For instance, *S. venezuelae* has a rifampin MIC of 0.5-1  $\mu\text{g}/\text{mL}$  but is perfectly capable of detoxifying rifamycins. If the true prevalence of rifamycin inactivators is 70-80%, it should be trivial to demonstrate this. We chose a small number of strains from the Wright Actinomycete Collection (~10 000 strains) at random and assayed them for inactivation. Rifamycin SV (5  $\mu\text{g}/\text{mL}$ ) or vehicle control (DMSO) were added to dense, 2-day-old liquid cultures of each WAC isolate. After 36 hours of incubation, culture supernatants were spotted on lawns of *Bacillus subtilis* to monitor inactivation (**Figure 7A**). We previously used this procedure to monitor inactivation in *S. venezuelae*, so this dosage should be appropriate for even relatively rifamycin-susceptible strains<sup>11</sup>. Of the 16 WAC isolates, 10 showed obvious inactivation (++ and + in **Figure 7B**), 3 showed no inactivation, and another 3 showed small/ambiguous drops in rifamycin activity (**Figure 7B**). Not considering the ambiguous cases, we observed inactivation in 10/16 isolates (62.5%), consistent with our genomic analyses and the notion that rifamycin inactivation is exceedingly common in soil actinomycetes. Curiously, many rifamycin inactivators appear to be comparatively susceptible to these compounds.



**Figure 7 Rifamycin inactivation is common in soil actinomycetes. A) Schematic for rifamycin inactivation assay.** The decline in Rifamycin SV (RifSV) concentration can be observed by adding culture supernatant to a lawn of susceptible bacteria, *Bacillus subtilis*. A rifamycin-resistant *B. subtilis* is used as a control to ensure the killing is RifSV dependent. **B) Inactivation of RifSV by random WAC isolates.** Inhibition of *B. subtilis* by WAC isolates incubated with 5  $\mu\text{g/mL}$  RifSV or DMSO for 36 hours. Shrinking or disappearing zones of inhibition indicate inactivation. Note WAC03356 is making an antibiotic in the absence of RifSV.

## DISCUSSION

Most attempts to catalog resistance genes, and study entire resistomes, rely strictly on protein-protein comparisons and use similarity at the amino acid or nucleotide level to infer equivalent function<sup>30</sup>. This approach works best for highly similar proteins but is less reliable for genes in distantly related organisms, even if these genes are functionally interchangeable. For instance, differentiating a rifamycin glycosyltransferase from a biosynthetic glycosyltransferase may be challenging *a priori*<sup>31</sup>. The regulation of a gene by the RAE belies its ultimate function – responding to rifamycins, presumably to confer resistance. This glimpse into the bacterial decision-making process removes the ambiguity of comparing distantly related proteins; for instance, we can confidently assume that a glycosyltransferase associated with the RAE acts on rifamycins. This study cataloged inducible rifamycin resistance across an entire bacterial phylum by exploiting a highly conserved regulatory feature, the extended RAE.

Using a curated database of representative Actinobacteria containing the highest quality genome for each species, we identified >3000 individual RAEs. The gene product controlled by each RAE was extracted and loosely grouped into protein families (**Figure 2**). We were unsurprised to find that known enzymes of the rifamycin resistome comprised major families (HelR, Rgt, Rox, Rph, Arr), but we were surprised to find that these 5 families comprised >97% of all RAGs<sup>10,11</sup>. We also considered the possibility that RAEs control operons of consecutive rifamycin resistance genes. Perhaps microbes acquire new resistance genes placed behind existing ones, allowing hosts to control additional mechanisms without losing existing ones or encoding more RAEs. We analyzed the next

two gene products downstream from the CDS directly controlled by the RAE (CDS1) (**Figure 1, Supplementary Figure 1**). We found 177 operons encoding *rgt-helR* and 18 instances of *rox* in the second position of a putative RAE operon, but these appeared to be the only abundant examples. Several small families, immediately downstream from a RAE or in the 2<sup>nd</sup>/3<sup>rd</sup> position, were also identified and contain proteins with plausible roles in rifamycin resistance; future work should attempt to characterize any role these may play in resistance (**Table 1, Supplementary Table 1**). This work suggests that we have found the major mechanisms of inducible rifamycin resistance. Furthermore, the combination of these 5 genes and *rpoB* polymorphisms may represent the entirety of the rifamycin resistome in Actinobacteria.

With the major families of RAGs identified, we next examined their taxonomic distribution and relative abundance. Firstly, the sheer abundance of the RAE was striking. From 3374 genomes, we identified 3328 RAEs, almost one per Actinobacteria. There is a general trend towards RAE abundance and the lifestyle of specific clades. For instance, taxa encompassing soil saprophytes and filamentous spore formers encode more RAEs (**Figures 3 and 4**). Furthermore, the RAE is missing from anaerobic taxa like Bifidobacteraceae and seems to have been lost more recently in the *Corynebacterium* (**Figures 3 and 4**). Apart from several deeply branching clades, the RAE can be found across most Actinobacterial families and is most commonly associated with HelR and Rox enzymes. In particular, the Streptosporangiaceae, Mycobacteraceae, Pseudonocardiaceae, Micromonosporaceae, and Streptomycetaceae are enriched in RAGs. Except for the Mycobacteraceae, these taxa contain the producers of all rifamycin antibiotics. The RAE

was not found in any rifamycin producers but was abundant in closely related species lacking rifamycin clusters, best exemplified by the *Amycolatopsis* (**Supplementary Figure 2**). The development and maintenance of resistance in organisms inhabiting the same ecological niche as rifamycin producers is an intuitive explanation for the abundance of RAGs in these taxa.

Of all RAGs, HelR and Rox are the most broadly distributed. In contrast, Rgt, Rph, and Arr are more prevalent in certain families. For instance, Rgt is rare outside the Sporangiaceae, Streptomycetaceae, and Pseudonocardaceae; the latter taxa encodes the most Rgt/genome by a significant margin. Similarly, Rph is present in most Micromonosporaceae species, and while it is found outside this family, it is far less abundant in other taxa. Lastly, Arr shows an extremely limited range, with >95% found within the Mycobacteriaceae. Arr must have been recently co-opted for rifamycin resistance by this clade. Additional analyses will be required to understand the evolution of inducible (RAE-associated) rifamycin resistance within Actinobacteria. For instance, the contribution of horizontal gene transfer to the spread of RAGs is unknown. Still, the distribution of specific resistance enzymes observed in this study suggests this trait is ancient. We speculate that it began with the emergence of the RAE to control the expression of HelR and Rox progenitors in an ancient Actinobacteria and that the recruitment of Rgt, Rph, and Arr enzymes likely came after and are derived from the Pseudonocardaceae, Micromonosporaceae, and Mycobacteriaceae respectively. We believe this dataset presents a unique opportunity for future studies to examine the role of regulatory capture alongside the evolution of dedicated resistance enzymes in the ancient history of antibiotic resistance.

For instance, do orthologs and paralogs of HelR and Rox which are not RAE associated retain the ability to confer resistance? Do novel resistance genes evolve the capability to confer resistance and then become tightly regulated by systems such as the RAE? Are genes driven towards a role in resistance due to fortuitous high levels of expression in the presence of a given antibiotic? Furthermore, once a regulatory system like the RAE exists, do recombination and duplication events that place new genes under its control accelerate the evolution and incorporation of new mechanisms of resistance?

This genomic-centric approach has revealed that RAE-associated rifamycin resistance is startlingly common among soil Actinobacteria. Approximately 80% of Streptomycetaceae encode RAEs, and ~70% encode rifamycin-inactivating enzymes with similar numbers from other common actinomycetes (**Supplementary Figures 3 and 4**). Rifamycins are not considered a common antibiotic to isolate during screening<sup>32</sup>. Still, they must be broadly distributed in nature and exert significant pressure on neighboring microbes to select for such widespread resistance. Bacteria that make rifamycins inhabit diverse soil and sediments, ranging from temperate soils<sup>33</sup>, not unlike those in your backyard, to the hyper-arid Atacama Desert<sup>34</sup> and even the ocean floor<sup>35</sup>. Furthermore, producers can be readily isolated in the lab using selective media containing rifampin<sup>36</sup>. The scars of the ancient and contemporary struggle to overcome rifamycins are visible in the genomes of Actinobacteria, a history we must consider when we re-appropriate these compounds for human use and mass production.

Confusingly, phenotypic rifampin resistance is much more uncommon than one would predict based on the presence of RAGs from our genomic data. We tested a small

sample (n = 16) of soil actinomycetes for rifamycin inactivation. We found that, in agreement with the genomic data, most of these bacteria could indeed inactivate the natural product rifamycin SV (**Figure 7**). We previously observed that for soil actinomycetes with a rifampin MIC >20 µg/mL, roughly half could inactivate the drug<sup>29</sup>. The genomic and phenotypic data generated here suggests that the >50% of soil actinomycetes are capable of inactivating rifamycins regardless of any resistance cutoff. This disconnect between the phenotypic susceptibility and the ability to degrade rifamycins is interesting. We know from *S. venezuelae*, which is rifampin susceptible, that knockout of RAGs does indeed increase the susceptibility of these organisms, and their association with the RAE implies that they function directly and intentionally in rifamycin resistance<sup>11</sup>. These aren't moonlighting enzymes with a fortuitous function. Our concept of meaningful resistance levels may be poorly calibrated to bacteria in their natural environment. The expression of these genes from the RAE may be insufficient for high-level resistance like we see during heterologous expression in *E. coli* but more than sufficient for life in the soil. Lastly, MIC assays mirror the human clinical use of antibiotics, whereby bacteria receive immediate exposure to a high drug concentration. In laboratory fermentations, bacteria synthesize antibiotics over 48-96 hours. Data on antibiotic production *in situ* is lacking, but it is reasonable to assume the kinetics are the same or even slower for bacteria in the soil. *In situ*, a rifamycin inactivator growing nearby a producing strain only needs to remove the antibiotic at a rate equal to its *de novo* production to stay below the MIC. The gradual, increasing exposure also gives rifamycin-inactivators time to reach maximal enzyme production, which they don't have in conventional susceptibility testing methods.

Conceptually, this may mean that ecologically relevant resistance (the ability to survive nearby an antibiotic producer) can go undetected in phenotypic screens for resistant environmental organisms, leading to an underestimation of the environmental resistome.

Mycobacteria are the predominant indication of rifamycin use in medicine, and we were interested in the presence of the RAE and associated resistance genes in *Mycobacterium* and related genera. Within the Mycobacteraceae, inducible rifamycin resistance varies profoundly between genera. *Corynebacterium* spp. has lost the RAE entirely, whereas the *Nocardia* universally possess multiple copies (**Figure 4**). The mycobacteria generally encode relatively few RAE-associated HelRs (**Figures 4 and 5**) compared to other families with similarly high RAE(s) carriage. This may result from some HelRs becoming de-coupled from the RAE. We observed in Chapter 4 that HelR's from mycobacteria, while still being highly similar, form a distinct cluster from other HelRs (also visible in the SSN in **Figure 2**). The proportion of HelR enzymes in this cluster associated with a RAE is far lower than the main group of HelRs, suggesting that some may be regulated in an alternative manner or perhaps do not confer rifamycin resistance<sup>11</sup>. However, we used different methodologies to detect RAEs in these two studies, so the results are not necessarily directly comparable. Beyond the specific issue of HelRs, it would be interesting for future work to investigate instances where RAE-associated genes exist de-coupled from the RAE. This would offer insights into their acquisition, evolution, and potentially their ancestral function. *Mycobacterium*, in the aggregate, possess at least one RAE in 76% of member genomes, but within the genera, a stark contrast exists between the slow-growing and fast-growing species (**Figure 5**). Fast growers are replete with RAEs

and the Arr enzyme specifically. The cause for this discrepancy could reflect differences in the ecology of these organisms and their predisposition to encounter rifamycins, which has encouraged the gradual loss of rifamycin resistance from the slow-growers.

This work highlights previously unidentified rifamycin resistance genes in many medically relevant species, such as *M. avium*, *M. intracellulare*, and *M. kansasii*, which all encode Rox enzymes (**Figure 5**). *In-vitro*, these non-tuberculosis mycobacteria (NTM) are considered rifamycin susceptible, and patients receive regimens containing rifamycins<sup>37,38</sup>. It is unknown if these enzymes contribute meaningfully to treatment failure in these species. Rifampicin doses may be sufficiently high to overcome the activity of these enzymes, or that induction proceeds too slowly to confer significant resistance, perhaps as a side effect of their slow growth. Conversely, it is possible that Rox-mediated resistance could play a more prominent role in rifamycin resistance or tolerance in the host. Previous work from our lab has shown that rifabutin cannot be inactivated by Rox<sup>39</sup>. Coincidentally this rifabutin is often used to treat these specific organisms. Generally, NTM infections are challenging to treat and feature low cure rates in part because of their high incidence in patients with HIV<sup>40</sup>. Rifabutin is favored over rifampin in these instances because it is a less potent inhibitor of host P450 enzymes making it easier to sustain concurrent antiretroviral therapy<sup>41,42</sup>. Several studies have shown better performance of rifabutin-containing regimens for MAC complex NTM<sup>41,43</sup>. Still, the complexity of the treatment for these cases makes it difficult to compare the specific antibacterial effects of rifabutin and rifampin. We also identify HelR, Rox, and Arr as the major mechanisms of intrinsic

rifamycin resistance across the fast-growing species (**Figure 5**), again offering insight into the design of next-generation rifamycins that can target these species.

*M. tuberculosis* belongs to a genus and family that possess a preponderance of RAGs and, luckily for us, is an outlier by not encoding any (**Figure 5**). From an evolutionary perspective, the parsimonious explanation is that *M. tuberculosis* lost rifamycin resistance in the distant past. As the ancestor of *M. tuberculosis* left the soil ecosystem and developed into an obligate pathogen of humans, rifamycin resistance was no longer required, and it may have lost this vestigial trait. In light of this possibility, our discovery and use of these compounds in the 20<sup>th</sup> century appear even more fortunate. Perhaps this general phenomenon played a role in antibiotics' initial, astounding efficacy against other common pathogens. On the other hand, emerging and opportunistic pathogens are often denizens of soil and water ecosystems, and in many instances, their full repertoire of intrinsic resistance remains intact.

Across all Actinobacteria, we detected approximately 1 RAE per genome. Taken at face value, this implies that most organisms encode a single RAE, but instead, many bacteria lack RAEs entirely, whereas many encode 2 or more (**Figure 6**). In fact, of organisms encoding at least one RAE, more possess  $\geq 2$  than 1. This trend is even more pronounced in specific taxa such as the Streptosporangiaceae, Streptomycetaceae, Mycobacteriaceae, Micromonosporaceae, and Pseudonocardiaceae (**Supplementary Figure 3**). Rather than combining RAGs randomly, we observed a consistent, logical pattern in bacteria with multiple RAEs. Organisms tend to encode a copy of HelR alongside 1-5 RAE-associated inactivating enzymes. As we show in Chapter 4, HelR can displace

rifamycins from RNAP, relieving their inhibition of transcription and releasing them into solution where they are made accessible to inactivating enzymes. Many Actinobacteria have harnessed this cooperative activity to provide maximal protection from rifamycins.

## **METHODS**

### **Construction of the extended RAE HMM**

We used BLASTn to search RefSeq whole genomes with the RAE as a query sequence with a 28.2 bit score cutoff<sup>44</sup>. After removing duplicates (hits to the same RAE in both orientations), we were left with 5169 sequences. Next, we used a python script to identify 724 RAEs that had a known rifamycin resistance gene immediately up or downstream by examining functional annotations for key terms ‘rifampicin’, ‘rifampin’, ‘phosphoenolpyruvate synthase’, ‘glycosyl transferase’, ‘phosphoenolpyruvate synthetase’, ‘FAD-dependent oxidoreductase’. This classification was not meant to be exhaustive but to generate a pool of RAEs that we could be confident were functional. Next, we pulled the sequence 100bp up and downstream from each RAE, and clustered them using usearchV10.0 (cluster\_fast command) at 100, 98, 95, and 90% identity and aligned the centroid sequences from each cluster<sup>45</sup>. We chose the alignment generated from the 95% identity clusters (596 sequences) based on manual inspection and trimmed this alignment to 37bp, which contained a high degree of conservation corresponding to the inverted repeat RAE sequence and the downstream -10. This alignment was used as a seed to generate the extended RAE HMM shown in **Figure 1**<sup>46</sup>.

### **Genomes used in this study and Identification of RAEs and associated genes**

All 3659 representative genomes from the phylum Actinobacteria (Actinomycetota) were downloaded from NCBI on October 5<sup>th</sup> 2022. We added *Bacillus subtilis* 168 as an outgroup for phylogenetic trees but didn't include this in any counting statistics. Additionally, we were curious about the phylogenetic distribution of rifamycin production and the presence of RAEs in producing organisms. Therefore, we added 15 genomes corresponding to known rifamycin producers, bringing the total number of genomes to 3374. Rifamycin producers included 7 from our collection (WAC07128, WAC09654, WAC10744, WAC02994, WAC03369, WAC06666, and WAC09165/9155 (see chapter 5)) and 8 from public repositories (Rifamycin B(GCF\_000220945.1), Rifamycin SV (GCF\_000384275.1), Rifamycin SV(GCF\_000425065.1), Kanglemycin A (GCF\_000716785.1), Chaxamycin(GCF\_001013905.1), Streptovaricin(GCF\_00870479 5.1), Rifamorpholines(GCF\_013364075.1), Rifamycin SV(GCF\_900091585.1)).

We determined optimal detection criteria using HMMscan (hmmer v3.3.2) on a small practice set of genomes and subsequently used those to identify RAEs in all 3374 genomes<sup>46</sup>. The optimal detection strategy was to use relatively loose inclusion criteria of an e-value of <0.01, in combination with a strict requirement for an alignment of  $\geq 30$ nt. We found that the e-value was significantly affected by the sequence of the spacer region, which shows low sequence conservation. The length criteria effectively required hits to span the inverted repeats and the -10 site on opposite sides of the model. This had the added benefit of suppressing the detection of each RAE twice, which occurs because the model can align to the inverted repeat portion in both directions. Still, the alignment length in the

incorrect orientation is invariably shorter, as it does not extend to the -10 site. These criteria identified a total of 3328 unique RAEs. Using a series of custom python scripts, we extracted the sequence 5kb up and downstream from each RAE, created a sequence feature corresponding to the RAE with the proper orientation, and pulled downstream CDSs (CDS1). As shown in **Figure 1**, instances where RAEs were upstream from pseudogenes, CDSs on the opposite strand (orphan RAEs), or if the contig ended before a CDS could be identified were counted separately. Successive CDSs, CDS2 and 3, were also included, provided they were in the correct orientation. In instances where CDS2 is on the opposite strand as the RAE, CDS3 was not included even if it was on the correct strand.

### **Sequence similarity network and categorization of RAE-associated genes.**

We constructed a sequence similarity network (SSN) to visualize and categorize the protein families associated with the RAE, using the same general strategy outlined in Copp *et al.* 2018<sup>47</sup>. An all-against-all BLAST search was performed on all RAE-associated proteins from the CDS1 position with an e-value cutoff of  $1 \times 10^{-22}$ , and the resulting SSN was visualized in Cytoscape (**Figure 2**)<sup>48</sup>. We constructed an independent SSN for the CDS2 and 3 genes using the same procedure (**Supplementary Figure 1**). From the CDS1 SSN, several major clusters corresponded to the known rifamycin-inactivating enzymes and HelR, in addition to a handful of smaller clusters and singletons, which we classified as miscellaneous RAGs. We extracted the nodes from each cluster and assigned each RAE in every genome an annotation (HelR, Rgt, Rox, Rph, Arr, miscellaneous, pseudogene, orphan, contig end). This genome-level inventory was used to generate the various counting, frequency, and proportional abundance statistics presented in this work.

### **Genome-based taxonomy and phylogenetic tree construction.**

All genomes were classified using the GTDBtk pipeline<sup>19</sup>. GTDB differs from NCBI in the placement of some taxa. For consistency, we used the GTDB classifications in all figures and analyses (**Figures 3-5, Supplementary Figures 2-4, Supplementary Table 2**). Notable differences were highlighted in the text wherever we deemed them relevant. GTDBtk extracts 120 single-copy essential genes from each genome, trims, concatenates, and aligns them in the process of classifying each bacteria. FastTree v2.1.11 was used to generate the phylogenetic tree from the multiple sequence alignment generated by GTDBtk using a WAG substitution model and default settings<sup>49</sup>. The Interactive Tree of Life (iTOL) platform was used to visualize trees and annotated features like RAEs, inactivating genes, rifamycin production, etc<sup>25</sup>.

### **Rifamycin inactivation.**

20 WAC isolates were chosen randomly from our in-house collection, with the only selection criteria being that they were not isolated on rifampin. Each strain was cultured first on Bennett's media and examined for purity. 3 isolates appeared to be fungi instead of actinomycetes, and a 4<sup>th</sup> strain was contaminated, so we omitted these isolates from the experiment. The 16 remaining strains were grown for 5 days in 3mL Streptomyces Antibiotic Medium at 30°C at 250rpm with sterile glass beads to homogenize the cultures. These starter cultures were diluted 1:20 into 2x3mL Tryptic Soy Broth (TSB) cultures and allowed to grow for 48 hours. Next, both cultures of each strain were centrifuged for 10 minutes at 4000 x g, and the cells were resuspended in 3mL of fresh TSB containing either 5 µg/mL Rifamycin SV (Sigma Aldrich) or DMSO as a vehicle control. After 36 hours of

incubation, 1 mL of methanol was added to each culture to solubilize the remaining rifamycin SV and cells were removed by centrifugation at 17 000 x g. 40µL of supernatant was applied to sterile cellulose discs placed on lawns of *Bacillus subtilis* 168 (rifamycin-susceptible) and *B. subtilis* 168 RpoBH482Y (rifamycin-resistant) on Tryptic Soy Agar plates<sup>31</sup>. After 16 hours of incubation at 37°C, the plates were examined for growth inhibition from the culture supernatants. A media control, which contained TSB + 5 µg/mL Rifamycin SV, was used to benchmark the level of inhibition in the absence of any inactivation.

## REFERENCES

1. O'Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. London: Review on Antimicrobial Resistance, 2016. doi:10.4103/2045-080x.186181
2. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
3. Wright, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* **5**, 175–186 (2007).
4. Waglechner, N., Elizabeth, C. J. & Wright, G. D. Ancient Antibiotics, Ancient Resistance. *EcoSal Plus* **9**, eESP-0027-02020 (2021).
5. *WHO Model List of Essential Medicines 19th List WHO Model List of Essential Medicines (April 2015) Explanatory notes.* (2015).
6. Floss, H. G. & Yu, T. W. Rifamycin Mode of Action, Resistance, and Biosynthesis. *Chem. rev.* **105**, 621–632 (2005).
7. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber. Lung Dis.* **79**, 3–29 (1998).
8. Campbell, E. A. *et al.* Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **104**, 901–912 (2001).
9. Goldstein, B. P. Resistance to rifampicin: a review. *J. Antibiot. (Tokyo).* **67**, 625–630 (2014).
10. Surette, M. D., Spanogiannopoulos, P. & Wright, G. D. The Enzymes of the Rifamycin Antibiotic Resistome. *Acc. Chem. Res.* **111**, (2021).
11. Surette, M. D., Waglechner, N., Koteva, K. & Wright, G. D. HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics. *Mol. Cell* **82**, 3151-3165.e9 (2022).
12. Spanogiannopoulos, P., Waglechner, N., Koteva, K. & Wright, G. D. A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci.*

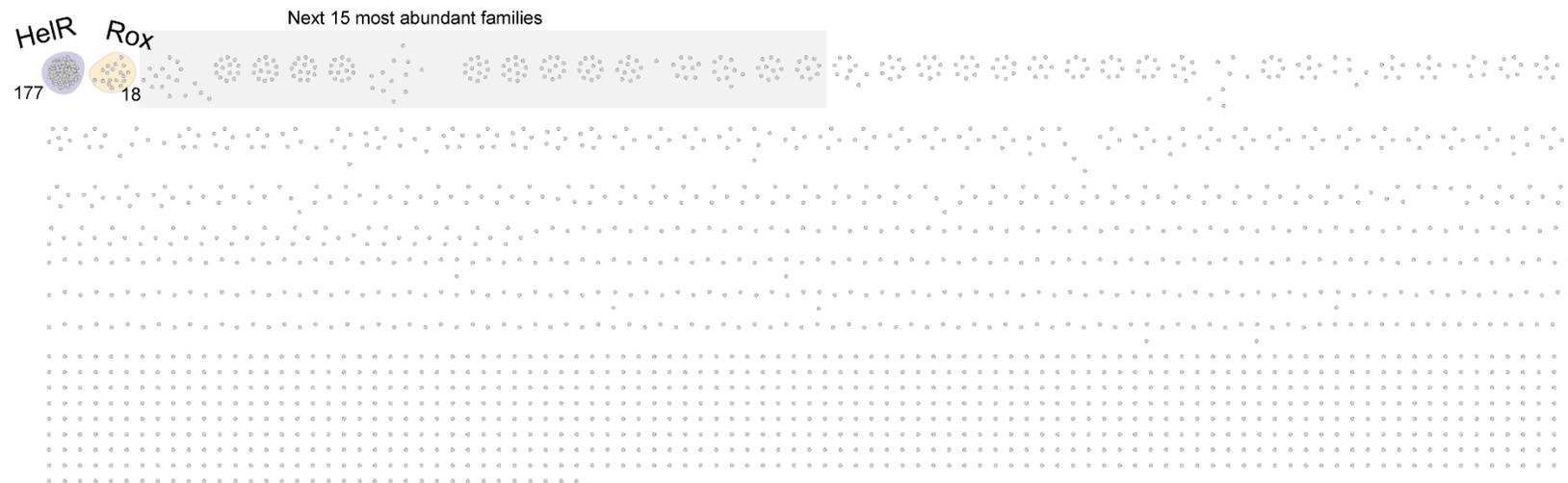
- U. S. A.* **111**, 7102–7107 (2014).
13. Shahmuradov, I. A., Mohamad Razali, R., Bougouffa, S., Radovanovic, A. & Bajic, V. B. BTSSfinder: A novel tool for the prediction of promoters in cyanobacteria and *Escherichia coli*. *Bioinformatics* **33**, 334–340 (2017).
  14. Chevez-Guardado, R. & Peña-Castillo, L. Promotech: a general tool for bacterial promoter recognition. *Genome Biol.* **22**, 1–16 (2021).
  15. O’Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
  16. Lu, S. *et al.* CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **48**, D265–D268 (2020).
  17. Wang, P., Bashiri, G., Gao, X., Sawaya, M. R. & Tang, Y. Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis. *J. Am. Chem. Soc.* **135**, 7138–7141 (2013).
  18. Mashalidis, E. H. *et al.* Rv2607 from *Mycobacterium tuberculosis* is a pyridoxine 5'-phosphate oxidase with unusual substrate specificity. *PLoS One* **6**, (2011).
  19. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**, 1925–1927 (2020).
  20. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).
  21. Chen, M. Y. *et al.* *Rubrobacter taiwanensis* sp. nov., a novel thermophilic, radiation-resistant species isolated from hot springs. *Int. J. Syst. Evol. Microbiol.* **54**, 1849–1855 (2004).
  22. Neuenschwander, S. M., Ghai, R., Pernthaler, J. & Salcher, M. M. Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. *ISME J.* **12**, 185–198 (2018).
  23. Diop, K., Bretelle, F., Fournier, P. E., Raoult, D. & Fenollar, F. “*Collinsella vaginalis*” sp. nov., a new bacterial species cultivated from human female genital tract. *Hum. Microbiome J.* **2**, 19–20 (2016).
  24. Mendling, W., Palmeira-de-Oliveira, A., Biber, S. & Prasauskas, V. An update on the role of *Atopobium vaginae* in bacterial vaginosis: what to consider when choosing a treatment? A mini review. *Arch. Gynecol. Obstet.* **300**, 1–6 (2019).
  25. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
  26. Ishikawa, J., Chiba, K., Kurita, H. & Satoh, H. Contribution of rpoB2 RNA polymerase beta subunit gene to rifampin resistance in *Nocardia* species. *Antimicrob. Agents Chemother.* **50**, 1342–1346 (2006).
  27. Meehan, C. J., Barco, R. A., Loh, Y. H. E., Cogneau, S. & Rigouts, L. Reconstituting the genus *Mycobacterium*. *Int. J. Syst. Evol. Microbiol.* **71**, (2021).
  28. Gupta, R. S., Lo, B. & Son, J. Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front. Microbiol.* **9**, 1–41 (2018).
  29. D’Costa, V. M., McGrann, K., Hughes, D. W. & Wright, G. D. Sampling the Antibiotic Resistome. *Science.* **311**, 374–378 (2006).

30. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **51**, D690–D699 (2023).
31. Spanogiannopoulos, P., Thaker, M., Koteva, K., Waglechner, N. & Wright, G. D. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob. Agents Chemother.* **56**, 5061–5069 (2012).
32. Baltz, R. H. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J. Ind. Microbiol. Biotechnol.* **33**, 507–513 (2006).
33. Sensi, P. History of the Development of Rifampin. *Rev. Infect. Dis.* **5**, S402–S406 (1983).
34. Rateb, M. E. *et al.* Chaxamycins A - D, bioactive ansamycins from a hyper-arid desert *Streptomyces* sp. *J. Nat. Prod.* **74**, 1491–1499 (2011).
35. Kim, T. K., Hewavitharana, A. K., Shaw, P. N. & Fuerst, J. A. Discovery of a New Source of Rifamycin Antibiotics in Marine Sponge Actinobacteria by Phylogenetic Prediction. *Appl. Environ. Microbiol.* **72**, 2118–2125 (2006).
36. Thaker, M. N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
37. Griffith, D. E. *et al.* An Official ATS / IDSA Statement : Diagnosis , Treatment , and Prevention of Nontuberculous Mycobacterial Diseases. *Am J Respir Crit Care Med Vol* **175**, 367–416 (2007).
38. Brown-elliott, B. A., Nash, K. A. & Wallace, R. J. Antimicrobial Susceptibility Testing, Drug Resistance Mechanisms, and Therapy of Infections with Nontuberculous Mycobacteria. *Clin. Microbiol. Rev.* **25**, 545–582 (2012).
39. Koteva, K. *et al.* Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem. Biol.* **25**, 403–412.e5 (2018).
40. Procop, G. W. HIV and mycobacteria. *Semin. Diagn. Pathol.* **34**, 332–339 (2017).
41. Crabol, Y., Catherinot, E., Veziris, N., Jullien, V. & Lortholary, O. Rifabutin: Where do we stand in 2016? *J. Antimicrob. Chemother.* **71**, 1759–1771 (2016).
42. Blaschke, T. F. & Skinner, Mi. H. The Clinical Pharmacokinetics of Rifabutin. *Clin. Infect. Dis.* **22**, 15–22 (1996).
43. Shafran, S. *et al.* A Comparison of Two Regimens for the Treatment of *Mycobacterium avium* Complex Bacteremia in AIDS. *N. Engl. J. Med.* **335**, 377–383 (1996).
44. Camacho, C. *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, 1–9 (2009).
45. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
46. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, (2011).
47. Copp, J. N., Akiva, E., Babbitt, P. C. & Tokuriki, N. Revealing Unexplored Sequence-Function Space Using Sequence Similarity Networks. *Biochemistry* **57**, 4651–4662 (2018).
48. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models. *Genome Res.* **13**, 426 (2003).
49. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, (2010).

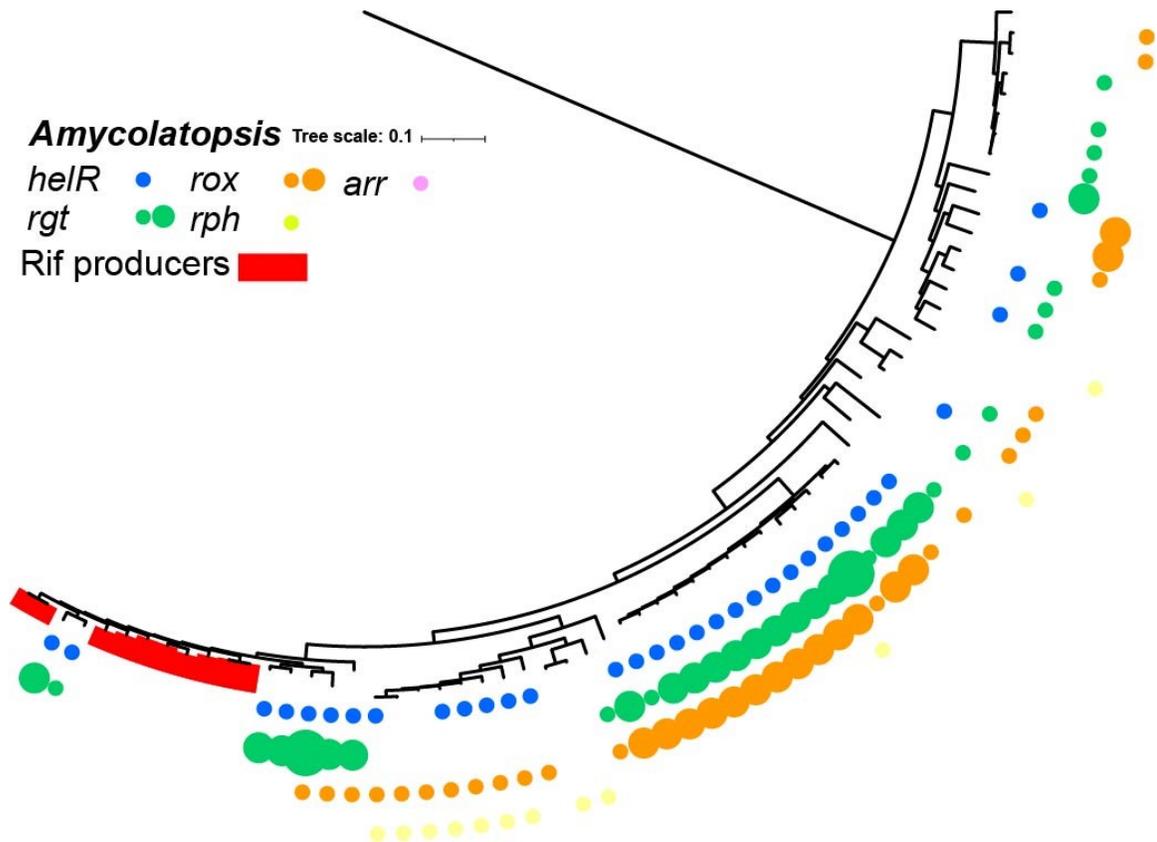
**Supplementary Material****Supplementary Table 1 – Miscellaneous proteins in CDS2 and 3 positions**

<b>Family</b>	<b>n</b>	<b>CDD*</b>	<b>Name</b>	<b>E-value</b>	<b>Function</b>
<b>M1-2</b>	16	CI32249	PRK07890 superfamily	3.21e <sup>-102</sup>	Short chain dehydrogenase
<b>M2-2</b>	13	CI35524	PRK06975 superfamily	3.95e <sup>-5</sup>	Uroporphyrinogen-III synthase/methyltransferase
<b>M3-2</b>	13	CI36191	fabG superfamily	1.94e <sup>-15</sup>	Ketoreductase
<b>M4-2</b>	13	PRK06185	PRK01685	0	FAD dependent oxidoreductase
<b>M5-2</b>	13	CI39137	YbjK superfamily	1.95e <sup>9</sup>	DNA binding regulator
<b>M6-2</b>	12	COG1131	CcmA	7.52e <sup>-68</sup>	Heme exporter ABC protein
<b>M7-2</b>	12	CI37187	PEP TPR lipo superfamily	1.49e <sup>-15</sup>	Secretion system-associated protein
<b>M8-2</b>	12	N/A	N/A	N/A	N/A
<b>M9-2</b>	12	N/A	N/A	N/A	N/A
<b>M10-2</b>	11	Pfam18029	Glyoxalase 6	1.28e <sup>-26</sup>	Glyoxalase
<b>M11-2</b>	11	Cd07247	SgaA N like	2.11e <sup>-27</sup>	N-terminus of A-factor binding protein
<b>M12-2</b>	10	Cd17321	MFS MMR MDR like	6.63e <sup>-61</sup>	Efflux pump
<b>M13-2</b>	10	Pfam07728	SpoIIE	1.50e <sup>-50</sup>	Involved in septum formation
<b>M14-2</b>	9	COG2197	CitB	3.34e <sup>-71</sup>	DNA binding response regulator
<b>M15-2</b>	9	Cd19088	AKR AKR13B1	7.38e <sup>-94</sup>	Phenylacetaldehyde dehydrogenase

\*Best hit from the conserved domain database (NCBI). N/A – no hit found



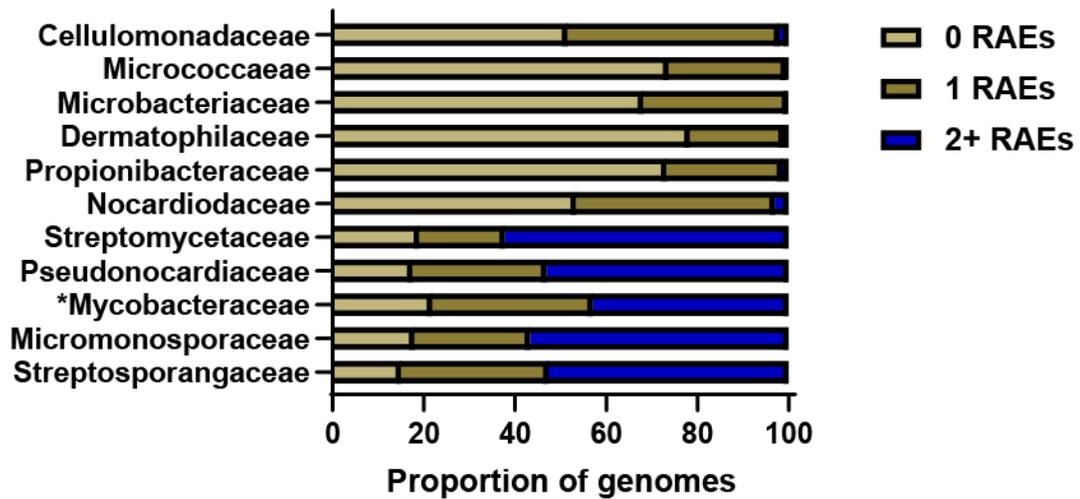
**Supplementary Figure 1 Sequence similarity network of RAE-associated CDS in the 2 and 3 positions.** All 2215 proteins in the 2<sup>nd</sup> and 3<sup>rd</sup> position in putative RAE operons clustered by sequence similarity. Protein clusters that belonged to a previously characterized family involved in rifamycin resistance were labelled accordingly. The preponderance of small clusters suggests that in most instances, the second and third genes downstream from the RAE are not actively controlled by this element.



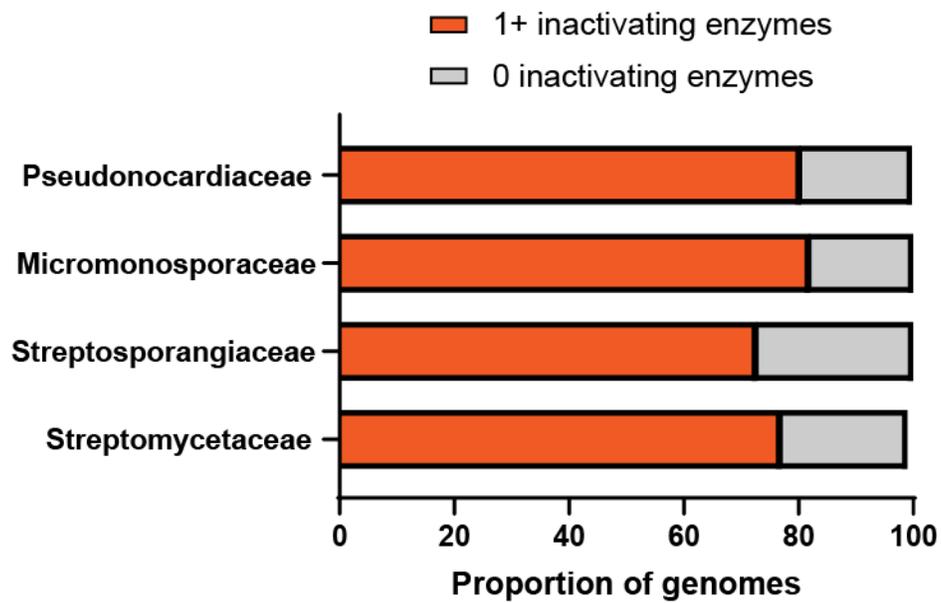
**Supplementary Figure 2 RAE inventory of *Amycolatopsis*.** Phylogenetic tree of the *Amycolatopsis* spp. from our dataset with RAE-associated resistance genes and rifamycin production indicated at the tips.

**Supplementary table 2 – Actinomycetaceae subdivisions**

Acintomycetaceae 1	Actinomycetaceae 2
<i>Georgenia A, Geogenia, Oceanitalia, Bogoriella</i>	<i>Bowdeniella, Flaviflexus, Peptidiphaga, Ancrocorticia, Neoactinobaculum, Actinobaculum, Actinotignum, Arcanobacterium, Arcanobacterium_A, Trueperella, Buchananella, Actinomyces, Mobiluncus, Winkia, UMGS822, Varibaculum, Boudabousia, Gleimia, Scrofinicrobium, Pauljensenia</i>



**Supplementary Figure 3 Encoding multiple RAEs is more frequent in specific taxonomic groups within Actinobacteria.** Proportion of genomes containing 0, 1, and >1 RAEs from all families encoding more than 50 members that encode at least one RAE. \*Mycobacteraceae stats have omitted the 140 *Corynebacterium* spp. which encode no RAEs.



**Supplementary Figure 4 Predicted frequency of rifamycin inactivating enzymes in select Actinobacterial families.**

**CHAPTER 4**

**HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics**

## PREFACE

The work presented in this chapter was previously published in:

Surette M.D., Waglechner N., Koteva K., and Wright G.D. 2022 HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics. *Mol. Cell.* 82(17):3151-3165. doi: 10.1016/j.molcel.2022.06.019.

Permission has been granted by the publisher to reproduce the material herein.

Waglechner N. performed HelD network analysis. Koteva K. synthesized and characterized RPP. M.D.S performed all other experiments. M.D.S. and G.D.W conceived of the experiments and wrote the manuscript.

## SUMMARY

Rifamycin antibiotics such as rifampin are potent inhibitors of prokaryotic RNA polymerase (RNAP) used to treat tuberculosis and other bacterial infections. While resistance arises in the clinic principally through mutations in RNAP, many bacteria possess highly specific enzyme-mediated resistance mechanisms that modify and inactivate rifamycins. The expression of these enzymes is controlled by a 19bp *cis*-acting rifamycin associated element (RAE). Guided by the presence of RAE sequences, we identify a helicase-like protein, HelR, in *Streptomyces venezuelae* that confers broad-spectrum rifamycin resistance. We show that HelR also promotes tolerance to rifamycins, enabling bacterial evasion of the toxic properties of these antibiotics. HelR forms a complex with RNAP and rescues transcription inhibition by displacing rifamycins from RNAP, thereby providing resistance by protecting RNAP. Furthermore, HelRs are broadly distributed in Actinobacteria, including several opportunistic *Mycobacterial* pathogens, offering yet another challenge for developing new rifamycin antibiotics.

## INTRODUCTION

Rifamycin antibiotics such as rifampin and rifabutin are semisynthetic derivatives of the natural product rifamycin B, discovered in 1957 as a fermentation product of *Amycolatopsis mediterranei* (Sensi, 1983). Rifampin (Rifampicin) emerged in the late 1960s and early 70's as a frontline treatment for infections caused by mycobacteria, particularly *Mycobacterium tuberculosis* (Floss and Yu, 2005). Rifamycins are potent inhibitors of prokaryotic DNA-dependent RNA polymerase (RNAP) (Wehrli et al., 1968). Bacterial RNAP is comprised of four proteins -  $\alpha_2$ ,  $\beta$ ,  $\beta'$ ,  $\omega$  - and a transiently associated  $\sigma$

factor (Chen et al., 2021). Rifamycins bind the  $\beta$  subunit of RNAP (RpoB) and occupy the path where the growing transcript emerges (Campbell et al., 2001). This interaction blocks the passage of newly synthesized RNA, impeding the production of mRNA longer than 2-3nt.

A pitfall of rifamycin antibiotics is their high frequency of resistance, approximately  $10^{-7} - 10^{-9}$  per bacterium per cell division (Gillespie, 2002). For this reason, they are primarily used in combination with other antibiotics. For example, against the slow-growing *M. tuberculosis*, rifampin is frequently combined with isoniazid and ethambutol (WHO, 2008). Genetic studies have identified an 81bp region in *rpoB*, termed the **Rifampin Resistance Determining Region (RRDR)**, accounting for ~95% of all rifamycin resistance mutations in the clinic (Yam et al., 2004). The spectrum of resistance mutations is relatively homogenous, with substitutions in three residues Asp516, His526, and Ser531 (*E. coli* RpoB numbering), which reduce the affinity of rifampin for the RNA exit tunnel and account for ~85% of rifampin resistant *M. tuberculosis* (Ramaswamy and Musser, 1998).

In contrast to the clinical resistome, the spectrum of rifamycin resistance mechanisms in the environment is highly diverse. While Gram-negative bacteria are primarily intrinsically insensitive due to impermeability and/or active efflux of these compounds, various Gram-positive genera have evolved highly specific mechanisms of rifamycin resistance. Many Actinobacteria, including but not limited to the genera *Streptomyces*, *Nocardia*, *Rhodococcus*, and many *Mycobacteria*, possess multiple mechanisms of enzymatic inactivation of rifamycins (Dabbs, 1987; Dabbs et al., 1995;

Spanogiannopoulos et al., 2012). Four distinct inactivation mechanisms are known: phosphorylation, ADP-ribosylation, glycosylation, and hydroxylation (**Figure 1**) (Dabbs et al., 1995; Koteva et al., 2018; Spanogiannopoulos et al., 2012, 2014; Surette et al., 2021). Arr enzymes catalyze ADP-ribose transfer to the C23 hydroxyl group (**Figure 1A**), thereby sterically blocking this essential hydroxyl required for binding to RpoB (Baysarowich et al., 2008). Rgt enzymes use UDP-glucose to glycosylate rifamycins at the same C23 hydroxyl group (**Figure 1A**) (Spanogiannopoulos et al., 2012). Rph enzymes transfer the  $\beta$ -phosphate from ATP to the adjacent C21 hydroxyl group (**Figure 1A**) (Spanogiannopoulos et al., 2014). This hydroxyl group also is essential for productive interaction with RpoB, and the addition of a bulky and negatively charged phosphate group at this position abolishes RNAP binding. On the other hand, the rifamycin monooxygenases (Rox) hydroxylate the naphthoquinone core and inactivate the drug by linearizing the rifamycin macrocycle, thereby destroying the three-dimensional structure of the antibiotic that is required for RNAP inhibition (**Figure 1A**) (Koteva et al., 2018).

Groups studying these inactivation mechanisms in the 1990s noted that enzyme production was often inducible by rifamycins (Quan et al., 1997). Over 20 years later, we discovered a 19 bp palindromic sequence upstream from *rgt*, which encodes a rifamycin glycosyltransferase (Spanogiannopoulos et al., 2014). This sequence, termed the Rifamycin Associated Element (RAE), was used to identify the gene encoding the enzyme responsible for rifampin phosphorylation in a different *Streptomyces* strain, revealing the RAE's predictive value in targeting rifamycin resistance genes. Subsequent bioinformatic analysis revealed that RAE sequences are found upstream of all known rifamycin

inactivating enzymes (Arr, Rox, Rph, and Rgt) within the Actinobacteria phylum. The RAE is found in many Actinobacterial pathogens such as *Mycobacterium abscessus*, *Rhodococcus equi*, and *Nocardia farcinia* but is absent from *M. tuberculosis*. The RAE was demonstrated to be necessary for inducing downstream genes in response to rifamycin antibiotics (Spanogiannopoulos et al., 2014). The molecular mechanism underlying this regulation remains unknown; nevertheless, the RAE, and the myriad of inactivating enzymes it controls, are a significant source of rifamycin resistance in Actinobacteria.

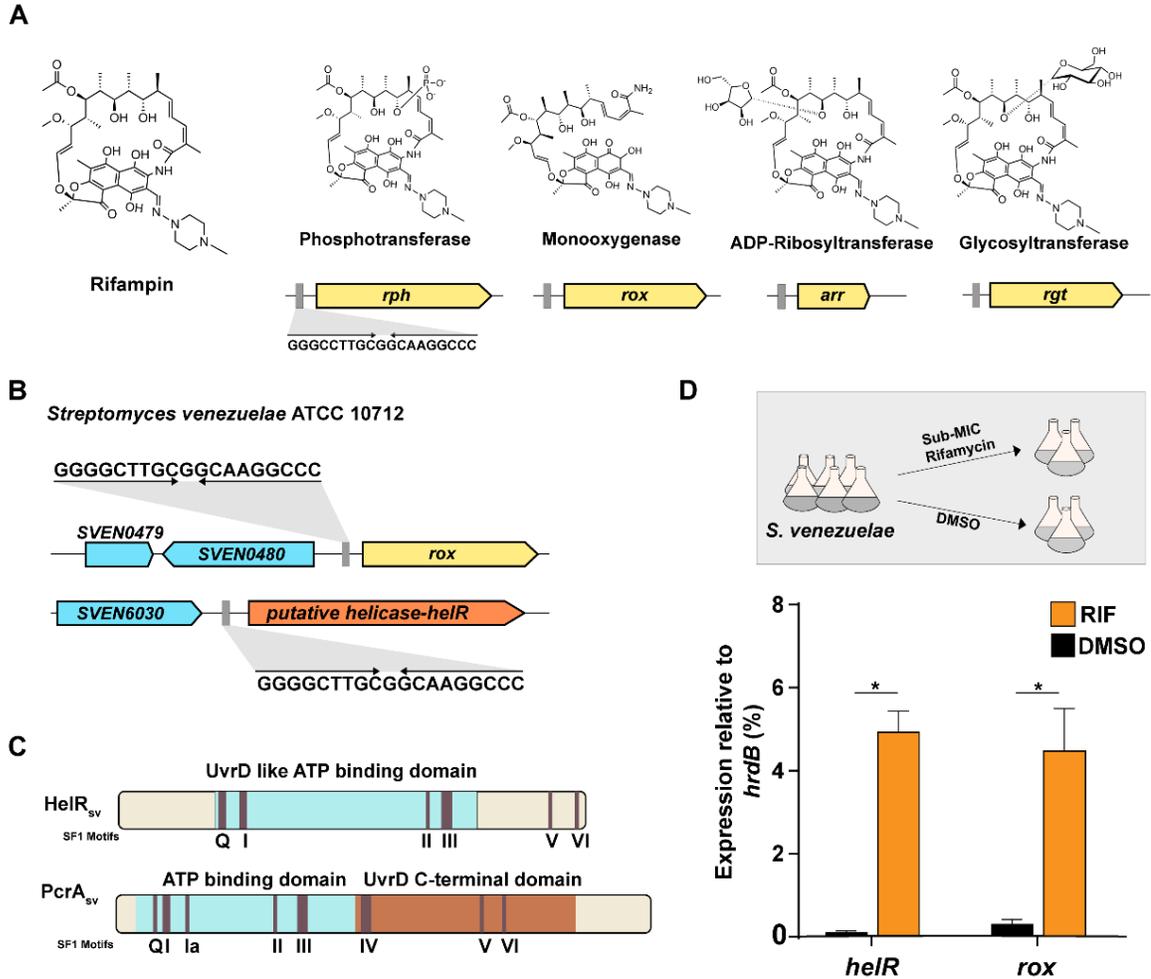
While all known rifamycin inactivation mechanisms appear associated with a RAE, the RAE is also found upstream of genes with no known resistance function. A significant proportion of RAEs are upstream of genes annotated as putative helicases (Spanogiannopoulos et al., 2014). Here we demonstrate that the rifamycin associated helicase-like protein (HelR<sub>sv</sub>) in *S. venezuelae* is a highly specific rifamycin resistance enzyme. In contrast to all previous RAE-associated genes, HelR is not an antibiotic inactivating enzyme. Instead, HelR directly interacts with RNAP, displacing bound rifamycins, thereby relieving inhibition. We also examine the distribution of HelR and the closely related HelD proteins, which are abundant in the genomes of Firmicutes and Actinobacteria. HelR homologs associated with RAEs fall into distinct protein clusters, indicating that only a small subset of these proteins are likely involved in rifamycin resistance.

## RESULTS

### *helR* is a rifamycin inducible resistance gene in *S. venezuelae*

In addition to its ubiquitous presence upstream of genes encoding rifamycin inactivating enzymes, the RAE is also associated with genes of unknown function. *Streptomyces venezuelae* ATCC10712 is a model *Streptomyces* with two RAEs in its genome (**Figure 1B**). One is associated with a rifamycin monooxygenase (Rox), and the other with a putative helicase (*helR*). HelR<sub>Sv</sub> is homologous to superfamily 1 helicases (**Figure 1C**), a protein class not previously associated with antibiotic resistance. Helicases are ssDNA translocases that couple ATP hydrolysis with directional movement along the DNA strand, where they collide with dsDNA and separate individual strands. While helicases are essential for DNA replication, they are not required for transcription (Dangkulwanich et al., 2014; Dillingham and Dillingham, 2011; Singleton et al., 2007). RNAP can melt promoter DNA to form a transcription bubble and open downstream DNA as the bubble moves during transcription without a helicase. Consequently, it was not apparent how the unwinding of any specific DNA or RNA segment could impact rifamycin activity, warranting genetic control by a RAE.

HelR is most similar to superfamily 1 helicases defined by core structural similarities and characteristic amino acid motifs (Fairman-Williams et al., 2010). We were able to identify many of these motifs in HelR<sub>Sv</sub> (**Figure 1C**). However, when we compared the domain architecture of HelR<sub>Sv</sub> to well-studied SF1 helicases like UvrD/PcrA using Interpro (Mitchell et al., 2019), we noticed that HelR lacks 1 of the two core SF1 helicase domains (**Figure 1C**). HelR is predicted to have a UvrD-like ATP binding domain that



**Figure 1 Rifamycin inactivation mechanisms and HelR are controlled by RAE sequences.** **A)** Structures of rifampin and its products inactivated by all known group transfer enzymes. The position of the RAE element is shown as a grey rectangle. **B)** RAE sequences are associated with *rox* and *helR* in *S. venezuelae*. **C)** Domain architecture of HelR and model Superfamily 1 helicase PcrA. Conserved Superfamily 1 motifs are depicted. **D)** Expression of *helR* and *rox* normalized to *hrdB* with and without exposure to sub-MIC rifamycins. P values calculated using an unpaired students t-test \* P < 0.001.

makes up the middle of the protein but lacks a UvrD-like C-terminal domain required for DNA binding. Without this domain, it is unclear how HelR could function as a helicase, and we hypothesized that it has a different function associated with rifamycin antibiotic activity.

We first confirmed the expected rifamycin-dependent expression of *helR* and *rox* using RT-qPCR (**Figure 1D**). RNA was isolated after a two-hour incubation with 0.5 µg/mL rifamycin SV (1/16X Minimal Inhibitory Concentration (MIC)) or DMSO (as vehicle control). The levels of *helR* and *rox* transcript were normalized to the constitutively expressed housekeeping gene *hrdB*. Both *helR* and *rox* showed low-level expression in the absence of rifamycin and were respectively induced 50-fold and 15-fold ( $P < 0.001$ ) in the presence of the antibiotic. Higher basal expression of *rox* appears to be the source of the difference in magnitude of induction, as both reach similar levels of maximal mRNA expression. Consistent with the presence of the RAE in its promoter region, *helR* is induced by rifamycins in *S. venezuelae*.

We generated single gene deletions of  $\Delta rox$  and  $\Delta helR$  as well as a double deletion of  $\Delta helR\Delta rox$  in *S. venezuelae* to explore the individual effects of both RAE-associated genes. HelR<sub>SV</sub> conferred robust resistance to all rifamycins tested (rifamycin SV, rifampin, rifabutin, and rifaximin), with increases in susceptibility ranging from 8- to 16-fold upon deletion (comparing  $\Delta rox$  to  $\Delta helR\Delta rox$ ) (**Table 1**). *S. venezuelae*  $\Delta rox$  was 4-8-fold more sensitive, except against rifabutin, an exceedingly poor substrate for Rox (Koteva et al., 2018). Surprised that *helR* confers higher resistance levels than *rox*, we confirmed that *helR* is not required for induction of *rox* (**Figure S1**). The MIC of rifampin for *S. venezuelae* falls >30-fold from an already low value of 0.5 µg/ml to 0.016 µg/mL in *S. venezuelae*  $\Delta helR\Delta rox$ , highlighting the efficacy of these dual resistance mechanisms. None of the strains tested showed altered susceptibility to fidaxomicin, an antibiotic that also targets RNAP but with a different binding site and mechanism of action than rifamycins (Lin et

al., 2018). Vancomycin and tetracycline, which respectively target cell wall biosynthesis and translation (Van Bambeke et al., 2004; Dürckheimer, 1975), also showed no difference in susceptibility. We could complement the *helR* mutant by cloning the gene into pIJ10257, an integrative vector that drives expression using the high-level constitutive promoter  $P_{ermE^*}$  (Hong et al., 2005). For all rifamycins, *S. venezuelae*  $\Delta helR\Delta rox$  pIJ:*helR* returned to the MIC levels of *S. venezuelae*  $\Delta rox$  and did not rise above it, indicating that overexpression of *helR* does not confer additional resistance.

The presence of conserved SF1 helicase motifs required for ATP binding and hydrolysis in HelR suggests that it has ATPase activity, which might be necessary for antibiotic resistance. Helicase motifs I and II (**Figure 1C**) correspond to the Walker A and B motifs found in many proteins that bind and hydrolyze ATP (Fairman-Williams et al., 2010). A conserved Asp and Glu in motif II are responsible for coordinating an essential catalytic  $Mg^{2+}$ , and the substitution of either of these amino acids abolishes ATPase activity in other helicases (Raney et al., 2013). We prepared an ATPase-impaired mutant by substituting the Asp of HelR's motif II (Val<sub>532</sub>AspGluAlaGln) to Ala. This mutant did not rescue resistance in *S. venezuelae*  $\Delta rox\Delta helR$  against all rifamycins tested (**Table 1**). In the case of rifampin, rifabutin, and rifaximin, HelR<sub>SvAsp533Ala</sub> lowered the MIC of *S. venezuelae*  $\Delta rox\Delta helR$  by two-fold rather than complement the loss of HelR<sub>Sv</sub>. The ATPase activity of HelR is therefore required for rifamycin resistance.

Having established that HelR is an ATP-dependent rifamycin resistance enzyme, we turned to elucidating its function. All previously characterized genes found under the

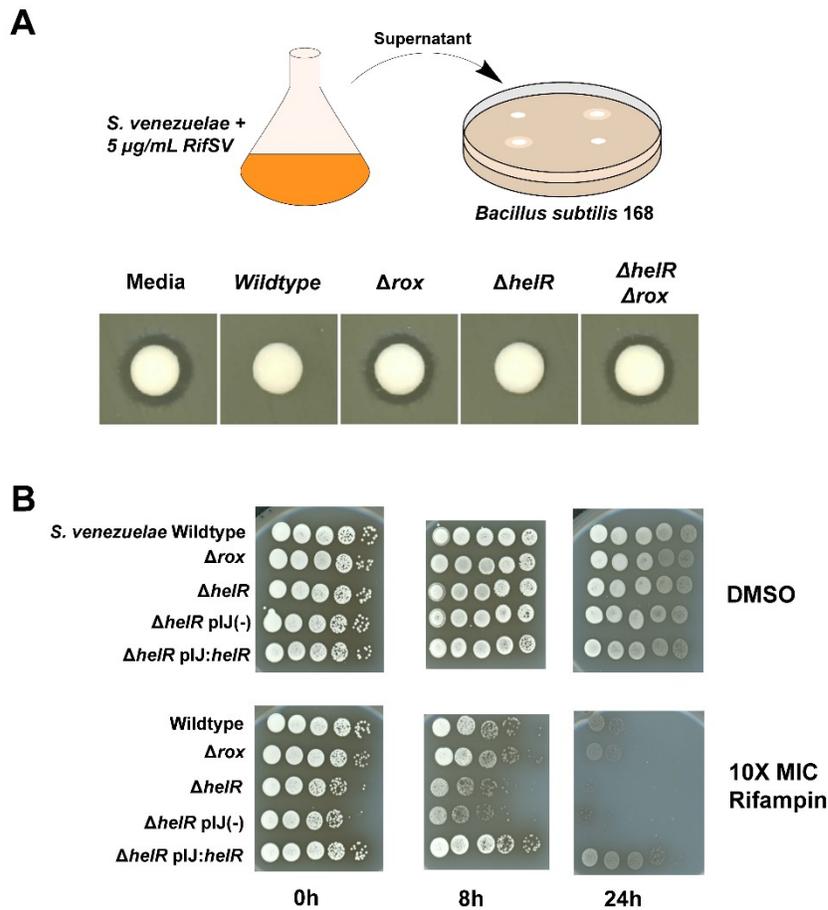
**Table 1 Susceptibility testing**

Strain	MIC (µg/mL)						
	Rifamycin SV	Rifampin	Rifabutin	Rifaximin	Fidaxomicin	Vancomycin	Tetracycline
<b><i>Streptomyces venezuelae</i> ATCC 10712</b>							
Wildtype	8	0.5	0.25	0.5	1	0.5	16
$\Delta$ rox	2	0.13	0.25	0.13	1	0.5	16
$\Delta$ helR	1	0.063	0.016	0.063	1	0.5	16
$\Delta$ rox $\Delta$ helR	0.13	0.016	0.016	0.016	1	0.5	16
$\Delta$ rox $\Delta$ helR pIJ10257(-)	0.13	0.016	0.016	0.016	1	0.5	16
$\Delta$ rox $\Delta$ helR pIJ:helR	2	0.13	0.25	0.13	1	0.5	16
$\Delta$ rox $\Delta$ helR pIJ:helR-Asp533Ala	0.13	<0.016	0.007	<0.016	1	0.5	16
$\Delta$ rox $\Delta$ helR pIJ:helR-FLAG		0.13					
$\Delta$ rox $\Delta$ helR pIJ:helR-his <sub>6</sub>		0.13					
$\Delta$ 5092		2					
$\Delta$ helR $\Delta$ 5092		0.063					
<b><i>Bacillus subtilis</i> 168</b>							
Wildtype		0.063					
$\Delta$ helD		0.063					

control of a RAE encode rifamycin inactivating enzymes, so we investigated this possibility for HelR. To monitor the inactivation of rifamycins, mycelia from 24-hour old cultures of *S. venezuelae* were washed and resuspended in fresh media containing 5 µg/ml of rifamycin SV and incubated for 24 hours to allow sufficient time for antibiotic inactivation. Culture supernatant was then applied to a lawn of rifamycin-sensitive *Bacillus subtilis* 168 to assess whether the antibiotic had been inactivated (**Figure 2A**). Antibiotic activity was lost from wild-type *S. venezuelae* and  $\Delta helR$  strains, consistent with inactivation of rifamycin SV due to the expression of *rox*. On the other hand, *S. venezuelae*  $\Delta rox$  and  $\Delta helR\Delta rox$  produced zones comparable to the media control (rifamycin SV without any *S. venezuelae*), indicating that Rox is the sole inactivating enzyme in *S. venezuelae* and that HelR must confer resistance through another mechanism.

### **HelR confers tolerance to rifamycins**

When conducting disc diffusion assays with rifamycins and *S. venezuelae* mutants, we noticed that if these plates were incubated for long periods of time, *helR*<sup>+</sup> strains would eventually grow within the initial zone of inhibition while *helR*<sup>-</sup> strains would not. We hypothesized that this might be due to the ability of HelR<sub>sv</sub> to confer tolerance, the ability to survive bactericidal concentrations of an antibiotic for more extended periods, in addition to resistance. Rifamycins are bactericidal towards Gram-positive bacteria and especially mycobacteria (Floss and Yu, 2005). We performed time-kill experiments on exponentially growing *S. venezuelae* with rifampin concentrations 10X the MIC of each strain. Cells were recovered and resuspended in fresh media three times to remove the remaining antibiotic, and tenfold dilutions were spotted onto agar to assess viability



**Figure 2A HelR is not an inactivating enzyme but confers tolerance to rifamycins. A)** *S. venezuelae* cultures were incubated for 24 hours in TSB + 5µg/mL rifamycin SV. The supernatant of this culture was applied to a cellulose disc and placed on a lawn of indicator bacteria (*B. subtilis*) to monitor antibiotic inactivation. Media control contained only TSB and rifamycin SV. **B)** *S. venezuelae* in exponential growth phase were standardized by OD<sub>600nm</sub> and exposed to 10X the strains respective MIC of rifampin. Cells were removed by centrifugation and washed several times to remove all remaining rifampin at the noted timepoints. Serial 10-fold dilutions were spotted onto Bennett’s agar to assess viability (dilutions are plated left to right).

(**Figure 2B**). Even at the 0 h timepoint where cells were only momentarily exposed to rifampin, *S. venezuelae*  $\Delta$ helR almost immediately lost ~10-fold viability. In contrast, wild-type *S. venezuelae*,  $\Delta$ rox, and  $\Delta$ helR pIJ:helR were unaffected. This difference became

more pronounced after 8 hours of antibiotic exposure. Relative to the starting inoculum,  $\Delta helR$  lost close to 100-fold viability, wildtype and  $\Delta rox$  show a very slight decline in viability, whereas  $\Delta helR$  pIJ:*helR* was fully viable. At 24 hours, a significant loss of viability for wild-type and  $\Delta rox$  cells is observed, although they are still orders of magnitude more viable than *helR*. At this timepoint,  $\Delta helR$  pIJ:*helR* is significantly more viable than wildtype cells. While overexpression of *helR* does not lead to a rise in MIC, it does increase drug tolerance. In addition to being an effective rifamycin resistance enzyme, these data show that HelR also allows *S. venezuelae* to tolerate inhibitory concentrations of these drugs for longer periods.

### **HelR forms a complex with RNA polymerase**

*Bacillus subtilis* HelD has the same overall domain architecture as HelR<sub>Sv</sub> consisting of a core UvrD-like ATP binding domain but lacking a C-terminal DNA binding domain, yet HelR<sub>Sv</sub> shares only minimal amino acid conservation with HelD (18% identity, 32% similarity). HelD lacks helicase activity *in vitro*, but it does bind RNAP (Wiedermannová et al., 2014). Furthermore, HelD stimulates transcription *in vitro* in an ATP-dependent manner (Wiedermannová et al., 2014). UvrD, a well-characterized and broadly conserved SF1 helicase, has also been reported to interact directly with RNAP during transcription-coupled repair (Epshtein et al., 2015). Based on this precedent, we hypothesized that HelR binds RNAP and that this interaction is linked to rifamycin resistance. We purified native RNAP from *S. venezuelae* grown in the presence of sub-MIC rifamycin (RIF+) and without (RIF-). We used the identical growth conditions used for the RT-qPCR experiments (**Figure 1D**) to ensure induction of *helR*, and we quantified

the proteins in each sample by LC MS-MS (**Figure 3A**). A total of 76 proteins were identified with high confidence, most of which did not change significantly in abundance between samples. HelR<sub>Sv</sub>, in contrast, was the most enriched protein (426-fold) following rifampin exposure (**Figure 3A, Supplemental File 1, Table S1**). Indeed, when purified RNAP fractions were analyzed by SDS-PAGE, a prominent band was visible at ~80kDa, consistent with the size of HelR<sub>Sv</sub>(**Figure 3B**). This band was excised from the gel and confirmed to be HelR<sub>Sv</sub> using LC MS-MS. We did not detect Rox in our samples, suggesting that the association of HelR<sub>Sv</sub> with RNAP is not an artifact of overexpression since this protein is also induced by rifamycin exposure.

Since purification of RNAP yields HelR<sub>Sv</sub>, we reasoned that the reciprocal experiment, purification of HelR<sub>Sv</sub> from *S. venezuelae*, should yield RNAP. We constructed a C-terminal FLAG-tagged HelR<sub>Sv</sub>, expressed this protein in *S. venezuelae*  $\Delta helR$ , and found that this construct restored wild-type levels of rifampin resistance and was therefore functional (**Table 1**). We performed co-immunoprecipitation using  $\alpha$ -FLAG resin from soluble proteomes of *S. venezuelae*  $\Delta helR$  constitutively expressing either native HelR<sub>Sv</sub> or HelR<sub>Sv</sub>-FLAG (**Figure 3C**). Western blotting confirmed the presence of HelR<sub>Sv</sub> and RNAP from cells expressing HelR<sub>Sv</sub>-FLAG and an absence of RNAP in cells expressing tag-free HelR<sub>Sv</sub>. These data show that RNAP and HelR<sub>Sv</sub> form a stable complex *in vivo*. Because HelR<sub>Sv</sub>-FLAG was expressed constitutively and co-precipitated with RNAP, the presence of rifamycin is not required for complex formation. The association of HelR<sub>Sv</sub> with RNAP in wildtype *S. venezuelae* is the consequence of rifamycin-mediated HelR<sub>Sv</sub> induction. The formation of a complex with HelR<sub>Sv</sub> and the molecular target of

rifamycin antibiotics lead us to hypothesize that HelR may function as a protection protein that prevents or reverses rifamycin binding in an ATP-dependent fashion (Wilson et al., 2020).

While writing this manuscript, we became aware of three co-structures of HelD in complex with RNAP published simultaneously; two describing HelD from *Bacillus subtilis* (HelD<sub>Bs</sub>) and the third from *Mycobacterium smegmatis*, which has 35% amino acid identity with HelR<sub>Sv</sub> (Kouba et al., 2020; Newing et al., 2020; Pei et al., 2020). The *M. smegmatis* protein is referred to as HelD<sub>Ms</sub> and is purported to be functionally equivalent to HelD<sub>Bs</sub>. However, we observed that *helD<sub>Ms</sub>* is associated with a RAE. Furthermore, this protein is among the most abundant in the cell following exposure to sub-MIC rifampin, and deletion of this gene results in increased susceptibility to rifampin (Hurst-Hess et al., 2019). We predict that this protein is functionally equivalent to HelR<sub>Sv</sub> and should be renamed HelR<sub>Ms</sub> to avoid confusion with genuine HelDs, which have no known role in inducible rifamycin resistance. Both HelD<sub>Bs</sub> and HelR<sub>Ms</sub> bind equivalent sites on RNAP, and both RNAP complexes appear incompatible with DNA in the primary channel. They both “wedge” open the  $\beta'$  Clamp weakening the interaction between RNAP and DNA and project appendages deep into RNAP. Additionally, both proteins possess a secondary channel arm (SCA) that is structurally similar to transcription factors such as GreA/B (Kouba et al., 2020; Pei et al., 2020). The SCA in HelD<sub>Bs</sub> extends into and occupies the active site. On the other hand, HelR<sub>Ms</sub> has a shorter SCA, which does not enter the active site; instead, this protein possesses a primary channel loop (PCh). Three distinct states of the HelR<sub>Ms</sub>-RNAP complex were solved by Kouba *et al.* The PCh loop's conformation was disordered in state

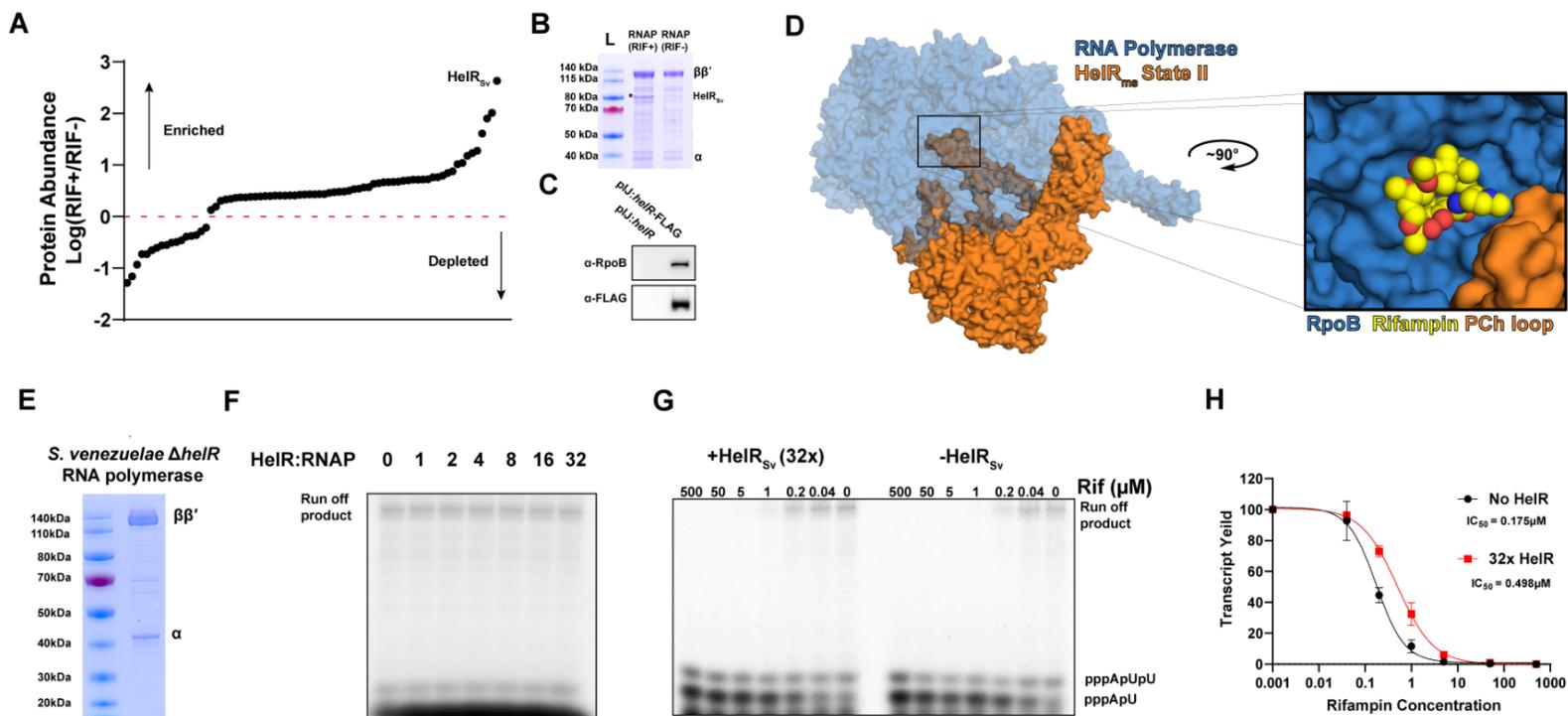
I, but in state II, it had folded into the primary channel and interacts with the catalytic  $Mg^{2+}$  and mobile domains in the active site. We noted that the path of the PCh loop comes very close to the rifamycin binding pocket on RpoB (**Figure 3D**). Our modeling suggests that the PCh loop does not clash with bound rifampin, which lies 3.5Å away (**Figure S2**), nor do key residues in this pocket appear to be significantly distorted in the structure of State II (**Figure S3**). Regardless, the existence of this appendage on a HelD-like enzyme that specifically confers rifamycin resistance is unlikely to be coincidental. Consistent with our hypothesis that HelR<sub>SV</sub> is an RNAP protection protein, HelR<sub>MS</sub> possesses an appendage uniquely suited for dislodging rifamycins from their target.

### **Mechanism of action of HelR**

To test our hypothesis that HelR is a rifamycin protection protein, we reconstituted transcription *in vitro* using RNAP purified from *S. venezuelae*  $\Delta helR$  to ensure that trace amounts of HelR<sub>SV</sub> did not contaminate preparations. To complete the reagent requirements for *in vitro* transcription, we purified the housekeeping sigma factor for *Streptomyces*  $\sigma^{HrdB}$  and synthesized the well studied promoter P<sub>AP3</sub> to serve as template (**Figure S4**). All attempts to express recombinant HelR<sub>SV</sub> with various affinity and solubility tags in *E. coli* were unsuccessful. To overcome this problem, we constructed HelR<sub>SV</sub> with a C-terminal His<sub>6</sub> tag and expressed this protein in *S. venezuelae*. As with the C-terminal FLAG tag, this construct completely restored rifampin resistance, indicating normal function (**Table 2**). We purified His<sub>6</sub>-tagged HelR<sub>SV</sub> directly from *S. venezuelae* for use in *in-vitro* assays.

Examination of HelD<sub>Bs</sub> and HelR<sub>MS</sub>:RNAP complexes suggest that binding of DNA and HelD are mutually exclusive, which means that RNAP can be engaged with DNA or HelD,

but not both. Somewhat paradoxically, HelD<sub>Bs</sub> can stimulate transcription in multiple round assays by removing stalled transcription complexes from template DNA. Interestingly HelR<sub>Ms</sub> has been shown to possess the ability to remove stalled elongating complexes, but there is no reported evidence that this protein can stimulate transcription *in vitro* (Kouba et al., 2020). We first added HelR<sub>Sv</sub> in increasing concentrations to determine if this protein had any general effects on *in vitro* transcription, taking care to operate at the linear portion of the reaction (**Figure 3F and Figure S4**). In multi-round assays, we found that HelR<sub>Sv</sub> had no discernable impact on the amount of transcript formed, even at a 32-fold molar excess over RNAP. The inability of HelR<sub>Sv</sub> to stimulate transcription in this context is further evidence that HelR is functionally distinct from HelD<sub>Bs</sub> and is instead a dedicated resistance enzyme. In agreement with our hypothesis that HelR functions to protect RNAP from rifamycins, we show that the addition of HelR<sub>Sv</sub> to *in vitro* transcription reactions offers protection from inhibitory concentrations of rifampin. A representative gel is shown in **Figure 3G**. The IC<sub>50</sub> of rifampin is raised from 0.175 μM (95%CI 0.142-0.214 μM) to 0.498 μM (95%CI 0.428-0.579 μM) in the presence of excess HelR<sub>Sv</sub>, a fold-change of 2.85 (**Figure 3H**).

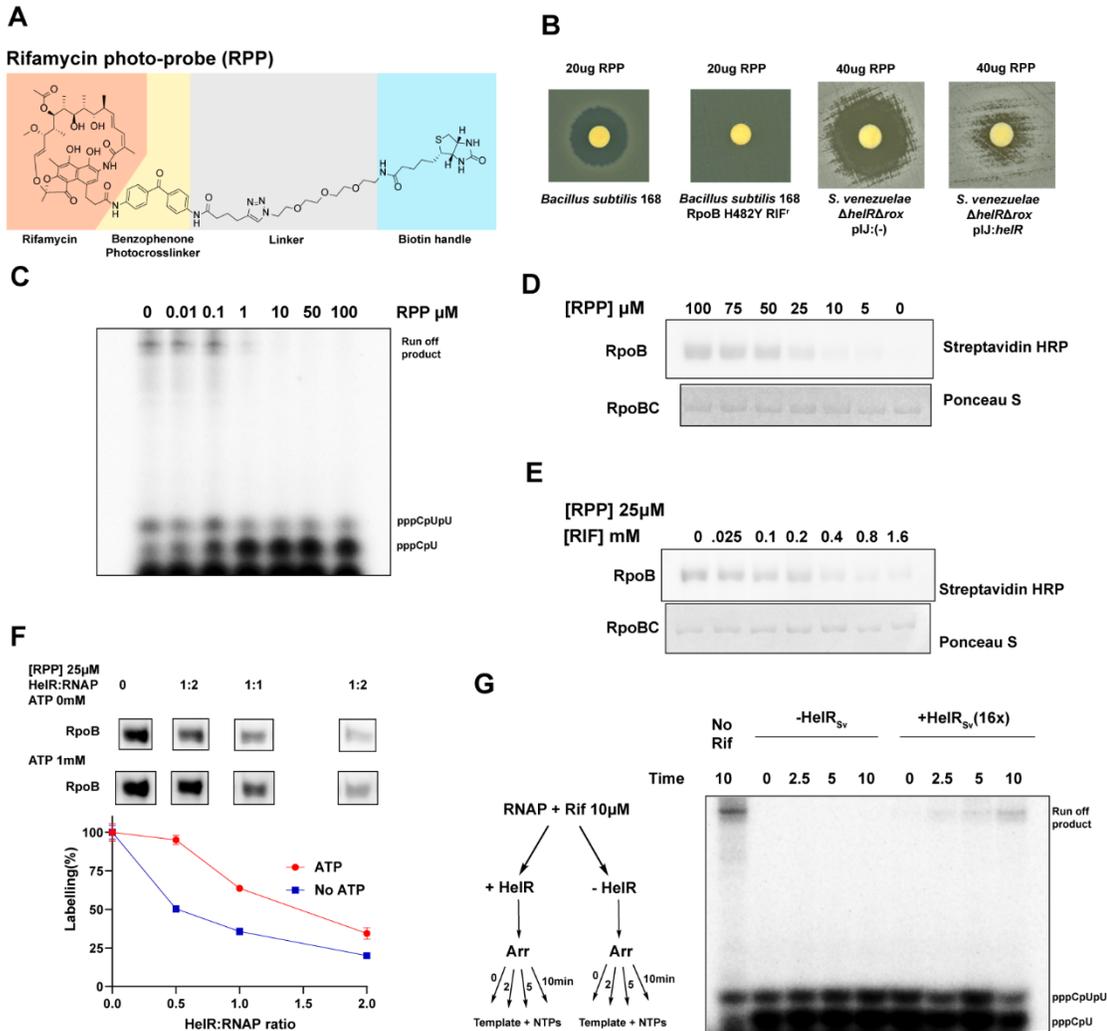


**Figure 3 HelR<sub>Sv</sub> co-purifies with RNA polymerase in *S. venezuelae* and blocks rifampin activity *in vitro*** **A)** Quantitative proteomic comparison of RNA polymerase purifications from rifamycin induced (RIF+) or non-induced (RIF-). Raw abundance data are expressed as the log of the ratio of RIF+ divided by RIF- and plotted in rank order. **B)** SDS-PAGE of RNAP preparations from induced and non-induced cells. The RNAP β, β' and α subunits are labeled. \*Denotes the band corresponding to HelR<sub>Sv</sub> **C)** Immunoprecipitation of soluble protein from *S. venezuelae* constitutively expressing native HelR<sub>Sv</sub> or HelR<sub>Sv</sub>-FLAG using α-FLAG resin. Protein was eluted from the resin using FLAG peptide and probed for the presence of RNAP (using an antibody which was raised against the β-subunit) and HelR (α-FLAG)

**(Figure 3) Continued D)** While engaging the main channel, the PCh loop of HelR<sub>Ms</sub> is in close proximity of the rifamycin binding pocket. Alignment of HelR<sub>Ms</sub>:RNAP complex from *M. smegmatis* (PDB ID 6YYS for State II) with a rifampin bound structure of *M. smegmatis* RNAP (6CCV) was used to model rifampin into the HelR<sub>Ms</sub>:RNAP complex **E)** SDS-PAGE of native RNA polymerase isolated from *S. venezuelae*  $\Delta helR$   $\beta$ ,  $\beta'$ , and  $\alpha$  subunits are labelled. **F)** Autoradiograph of transcripts from multiple round *in-vitro* transcription reaction with increasing molar ratio of HelR<sub>Sv</sub>:RNAP( $\sigma^{hrdB}$ ) using P<sub>AP3</sub> as template **G)** Representative gel of multiple round transcription reactions performed with rifampin (RIF) in the presence and absence of a large molar excess of HelR<sub>Sv</sub>. Abortive products generated by rifampin pppGpU and pppGpUpU from P<sub>AP3</sub> are labeled. **H)** IC<sub>50</sub> of rifampin in the presence or absence of a 32-fold molar excess of HelR<sub>Sv</sub>. Run-off transcripts were quantified using densitometry; data were generated from three independent experiments.

To better understand the molecular mechanism of HelR-mediated rifamycin resistance, we sought a more direct method to measure the presence of rifamycins bound to RNAP. Consequently, we designed and synthesized a rifamycin photoaffinity-probe (RPP) (**Figure 4A**). We used rifamycin B as a scaffold to take advantage of its free carboxylic acid in a location on the antibiotic known to be tolerant to substitutions. Using standard methods, we coupled a benzophenone-alkyne to rifamycin B through the carboxylic acid. Excitation of the benzophenone by long wavelength UV light (365 nm) generates a highly reactive carbene that can crosslink to nearby proteins (Murale et al., 2017). Lastly, we used ‘click chemistry’ to link the rifamycin-benzophenone to commercially available biotin-PEG3-azide, allowing us to detect the presence of the crosslinked probe using a streptavidin-HRP conjugate (Kolb et al., 2001). RPP recapitulates all relevant properties of a rifamycin antibiotic. It inhibits the growth of both *B. subtilis* 168 and *Streptomyces venezuelae*; a known rifamycin resistance substitution in RpoB (*B. subtilis* numbering, H482Y) confers RPP resistance; and importantly, *helR* also

provides increased resistance to RPP (**Figure 4B**). RPP inhibits RNAP *in vitro* at concentrations comparable to rifampin in multiple round assays and, like all rifamycins, results in the accumulation of abortive products (**Figure 4C**). We used RPP to photolabel purified *S. venezuelae* RNAP *in vitro*. Samples were exposed to increasing concentrations of RPP, proteins were separated by SDS-PAGE, transferred to PVDF membranes, and probed with streptavidin-HRP (**Figure 4D**). Blotting of purified RNAP revealed two endogenously biotinylated proteins present in our preparations; this had no impact on our ability to label and quantify the labeling of RNAP with RPP (**Figure S5A**). We observed concentration-dependent labeling of RpoB, which began to saturate at ~50  $\mu\text{M}$  RPP. At concentrations used for labeling (25  $\mu\text{M}$ ), crosslinking in the absence of UV excitation is negligible (**Figure S5B**). Addition of rifampin blocked the labeling of RNAP by RPP, again demonstrating that our probe is specific for the rifamycin-binding pocket (**Figure 4E, S5C**). If our hypothesis is correct and HelR<sub>SV</sub> can block/displace rifamycins, it should also affect labeling by RPP. We exposed RNAP to a gradient of HelR<sub>SV</sub> from 1:2 to 2:1 (HelR<sub>SV</sub> to RNAP) in the presence and absence of 1 mM ATP (**Figure 4F**). Labeling of RpoB declines as HelR<sub>SV</sub> concentration rises; at a 2:1 ratio of HelR<sub>SV</sub> labeling is decreased by 65.6% (95% CI 61.4 – 69.7) with ATP and 80.0 % without (95% CI 78.4-81.5). The addition of ATP lead to an increase in labeling relative to the no ATP reactions across all concentrations. We next used an orthogonal approach to demonstrate that HelR<sub>SV</sub> can displace rifampin from RNAP. We first formed complexes with RNAP( $\sigma^{\text{HrdB}}$ ) and rifampin by adding a large excess of the drug (10  $\mu\text{M}$ , >20x the IC<sub>50</sub>). We then added rifamycin



**Figure 4 Design and activity of a rifamycin photoaffinity probe (RPP) demonstrate that HelR<sub>sv</sub> displaces RNAP-bound rifamycins.** **A)** Structure of RPP with relevant elements highlighted. Rifamycin B scaffold (orange), Benzophenone (yellow), and Biotin (light blue) with the linker region in grey. **B)** Zones of growth inhibition resulting from RPP when spotted on a lawn of *B. subtilis* 168 wildtype and a rifampin resistant mutant (left) and *S. venezuelae* Δ*helR*Δ*rox* with empty vector or constitutive expression of *helR*. **C)** Autoradiograph of multiple round *in-vitro* transcription reactions in the presence of increasing concentrations of RPP **D)** RPP labels RNAP *in-vitro*. RNAP was incubated with an increasing gradient of RPP, and crosslinks were formed by exposure to 365nm light. Proteins were analyzed by Western blot using Streptavidin-HRP to detect covalently linked RPP, followed by Ponceau S staining for total protein. **E)** Competition between Rifampin and RPP. Rifampin was added at various concentrations before the addition of RPP (kept constant at 25 μM). **F)** HelR<sub>sv</sub> displaces RPP from RNAP. 25 μM RPP was used to label RNAP, which had been pre-incubated with HelR<sub>sv</sub> at various molar ratios, both in the presence and absence of 1 mM ATP. All reactions were performed in triplicate and quantified using densitometry, Error bars represent SD, all reactions are normalized to

**(Figure 4 Continued)** reactions with no HelR<sub>Sv</sub> (100% labelling). **G)** HelR displaces rifampin from RNAP *in-vitro*. RNAP was incubated in a saturating level of rifampin (>20X the IC<sub>50</sub>) and treated with the rifampin inactivating enzyme Arr alone or with Arr and HelR<sub>Sv</sub>. Control reaction without rifampin also contained Arr and HelR<sub>Sv</sub>. Transcription was then initiated at several time points by adding P<sub>AP3</sub> and NTPs; the products were analyzed by Urea-PAGE and autoradiography.

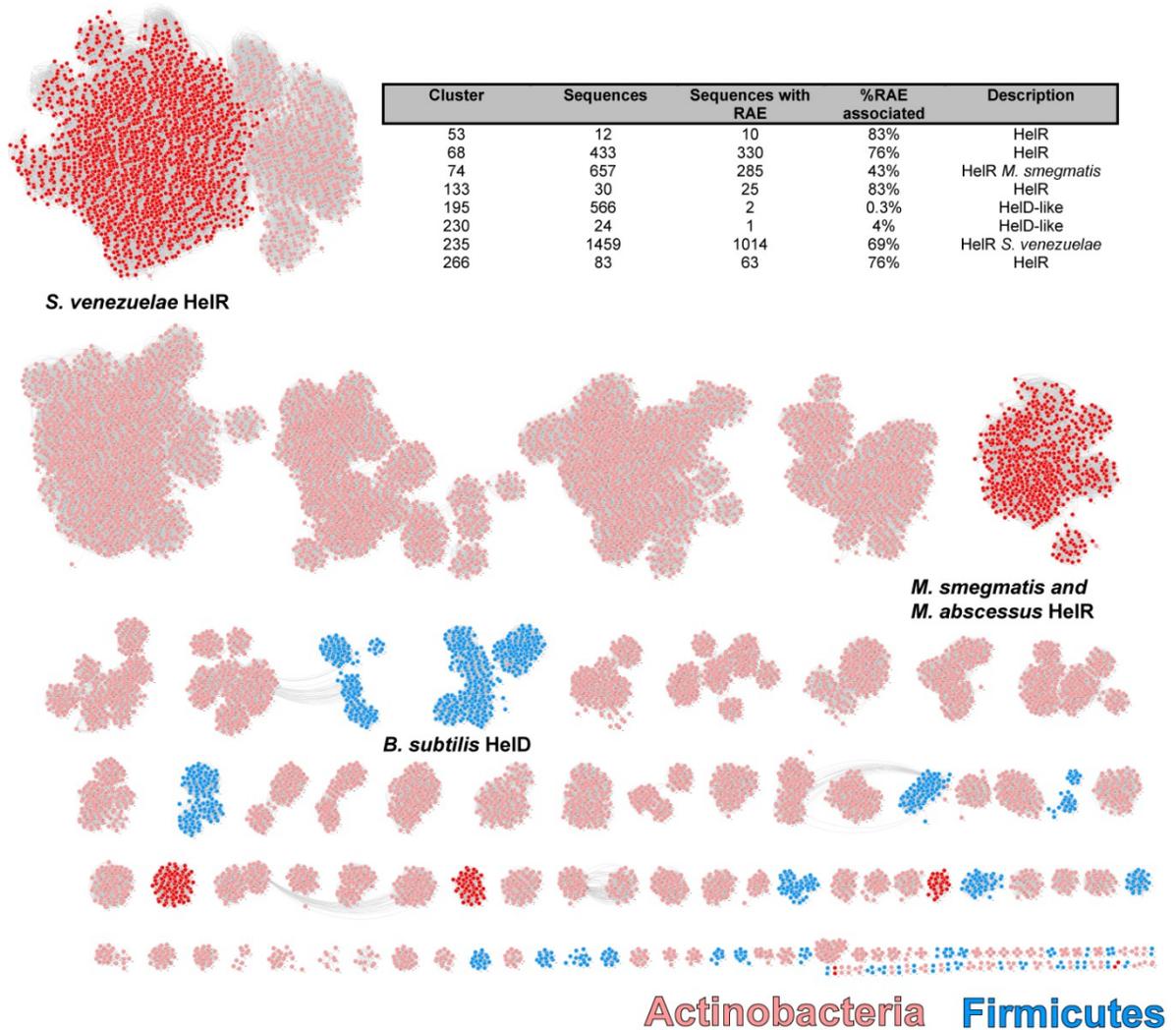
ADP-ribosyltransferase (Arr), which inactivates rifampin rendering it unable to bind (or rebind) RNAP. Since Arr only inactivates rifampin in solution and cannot access rifampin bound to RNAP, Arr cannot ‘rescue’ these polymerases independently and relies on the dissociation of rifampin from RNAP. We reasoned that the addition of HelR<sub>Sv</sub> could accelerate this process, and indeed we observe a time-dependent easing of transcription inhibition (production of full-length transcript) when Rifampin saturated RNAP is incubated with both Arr and HelR<sub>Sv</sub> but no such effect in the absence of HelR<sub>Sv</sub> (**Figure 4G**). These data demonstrate that HelR<sub>Sv</sub> functions as an RNAP protection protein by dislodging rifamycins bound to the enzyme.

### **HelD-like proteins in microbial genomes**

HelR belongs to a family of proteins not previously associated with antibiotic resistance. We surveyed the diversity of HelD-like proteins in bacteria and used the presence of a RAE to identify proteins likely involved in rifamycin resistance (HelRs). We ran BLASTp using HelR<sub>Sv</sub> and HelD<sub>Bs</sub> as queries and clustered the top 5000 hits at 99% identity, producing 4906 and 3026 representative sequences for HelR and HelD, respectively. As the non-redundant database includes sequences with varying levels of

annotation, we used these sequences as a large BLASTp query set against consistently annotated genomes in the RefSeq database. Using the representative non-redundant HelD/HelR protein sequences to query RefSeq Actinobacteria and Firmicute genomes returned 15,136 putative HelD/HelR sequences. At a 50% identity threshold, the sequences can be assembled into 417 cluster families. We removed several spurious clusters of glycosyltransferases (see **Supplementary Table 2**, and explanation in STAR methods) and then identified protein clusters with at least one member associated with a RAE. Only 29 of these cluster families, encompassing 3314 HelD-like sequences (21.9% of the total), contain at least one member associated with a RAE, and none of these originate from Firmicute genomes. 1774 protein sequences (out of 3314 or 53.5% of the sequences in the 29 RAE-associated clusters, 11.7% of the total putative HelD/HelR sequences) are associated with a RAE. 12 of these 29 clusters have only one member, and 21 clusters had <8 members, **Figure 5** shows a summary of the clusters with >8 members. We designated proteins as putative HelRs if they fell into a cluster with a high proportion of RAE association. The *S. venezuelae* and mycobacterial HelRs fall into the two largest clusters (235 and 74, respectively). Except for clusters 195 and 230, most proteins belonging to one of these clusters are associated with a RAE. This was notably lower in cluster 74, which encompasses the *Mycobacterial* sequences. We attempted to assemble all the HelD-like proteins into a comprehensive phylogeny to determine where the rifamycin resistance proteins fell on the tree, but these sequences proved too diverse. For instance, just the RAE-associated clusters 53, 68, 74, 124, 133, 195, 230, 235, and 266 produced an alignment with over 4000 sites (more than 4x the length of an individual HelD-like protein) and many

gaps. Automated trimming reduces this alignment to 357 sites with visually poor alignment. We concluded that it was impossible to produce a reliable phylogeny using all these sequences; they are too diverse for the usual approaches. An alternative method to represent the relationships between these sequences is a sequence similarity network (SSN)(Copp et al., 2018). We constructed an SSN from an all-against-all similarity search using an e-value cutoff of  $1e^{-10}$ . This network consists of 15137 nodes and 259,167 undirected edges after removing self-edges. We visualize this network in Cytoscape (Shannon et al., 2003) using the organic layout and include the cluster data from **Supplementary Table 2 (Figure 5)**. This network shows two distinct large clusters of HelRs (Clusters 235, 266, 68, 53) & (74, 230) and three small clusters. One consists of all the members of 133, and two clusters are made up of subsets of cluster 235 (**Figure S6**). We found no instances of Firmicute and Actinobacterial HelD-like proteins within a cluster. HelRs make up a distinct subset of all HelD-like enzymes, suggesting that most HelDs are not involved in rifamycin resistance and may have analogous biological functions to *B. subtilis* HelD or perhaps novel ones. We also noted an abundance of HelD-like proteins in *Streptomyces* genomes. This is not solely due to the overrepresentation of *Streptomyces* genomes in Refseq and reflects numerous copies of HelD-like genes per genome. For instance, *S. venezuelae* ATCC 10712 encodes 5 HelD-like proteins (including HelR<sub>Sv</sub>), whereas *M. smegmatis* and *M. abscessus* have only HelR and no other HelD-like proteins. HelR is also not strictly conserved among mycobacteria and is absent from medically important slow-growing species such as *M. tuberculosis*, *M. kansasii*, *M. ulcerans*, *M. leprae*, and *M. bovis*.



**Figure 5. Sequence Similarity network of HelD-like proteins** A network of 15,137 HelD-like proteins from Refseq genomes of Actinobacteria and Firmicutes. Each protein is represented as a node, color-coded by the phyla it comes from (light red for Actinobacteria, blue for Firmicutes) and whether it belongs to a RAE-associated protein cluster (Red). Cluster 195 was deliberately not colored red because it contains few RAE-associated genes. HelRs (RAE-associated HelD-like proteins) form two major clusters and three small ones, indicating most HelD-like proteins are not rifamycin resistance enzymes and likely have other functions.

We were intrigued by cluster 195, which has 566 members, of which only two are associated with a RAE (**Figure 5**). *S. venezuelae* encodes a member of cluster 195, SVEN5092, which we identified in our RNAP preparations, evidence that this protein also interacts with RNAP. Unlike HelR<sub>sv</sub>, SVEN5092's abundance was unchanged between rifamycin-induced and uninduced samples. We deleted this gene and were surprised to find that *S. venezuelae*  $\Delta$ 5092 had become 4-fold *more* resistant to rifampin. When we performed the same deletion in a  $\Delta$ helR background, resistance to rifampin was unchanged, suggesting that SVEN5092 may compete with HelR<sub>sv</sub> for RNAP binding. Nevertheless, SVEN5092 is an example of a HelD-like protein that does not confer rifamycin resistance. Curiously, even though it lacks a RAE, *helD* has been reported to be inducible by transcription inhibitors such as rifampin in *B. subtilis* (Hutter et al., 2004). This induction has been used as a reporter to identify the mechanism of action of novel antimicrobials (Mosaei et al., 2018). We confirmed that *B. subtilis*  $\Delta$ helD is no more susceptible to rifampin than the parent strain. HelRs are a minority of HelD-like enzymes and are unique in their ability to confer rifamycin resistance.

## DISCUSSION

The environmental resistome is the source of many antibiotic resistance elements that emerge in pathogenic bacteria (Surette and Wright, 2017). A complete understanding of the mechanistic diversity of antibiotic resistance is needed to preserve antibiotics in the resistance era (Brown and Wright, 2016). Characterizing these mechanisms can help us anticipate resistance before it emerges in the clinic, identify genes responsible for intrinsic resistance in pathogens, and guide the synthesis of new drugs. The capacity of

bacteria to engage in horizontal gene transfer means that resistance in the environment remains an existential threat to the continued clinical efficacy of antibiotics over the long term.

In this work, we use the presence of a *cis*-regulatory genetic element that specifically induces genes in response to rifamycin antibiotics to guide the discovery of a previously uncharacterized resistance gene, *helR*. Using *S. venezuelae* as a model system, we confirm that the expression of *helR* is rifamycin inducible and that *helR* confers resistance to a variety of rifamycins, both natural and semisynthetic drugs in current use.

Based on a shared domain architecture with the known RNAP interacting protein HelD, we hypothesized and verified that HelR<sub>sv</sub> directly interacts with RNAP. HelD-like enzymes, HelR included, share the core ATPase machinery of Superfamily 1 helicases. These proteins use the energy released from ATP hydrolysis to generate mechanical force to promote translocation along ssDNA and strand displacement. In contrast to true helicases, HelD-like enzymes lack DNA binding domains and instead use this mechanical force to remodel themselves and/or their interaction partner RNAP (Kouba et al., 2020; Newing et al., 2020; Pei et al., 2020). We demonstrated that ATP hydrolysis is essential for the resistance activity of HelR<sub>sv</sub> but that HelR<sub>sv</sub> does not inactivate the antibiotic. This observation is analogous to another group of antibiotic resistance enzymes, the ribosomal protection proteins such as TetO/TetM. These proteins are GTPases that bind to the ribosome and directly displace target-bound tetracyclines, allowing for resumption of translation (Dönhöfer et al., 2012; Li et al., 2013). HelR<sub>sv</sub> is an ATPase which binds to RNAP, so we hypothesized that it directly displaces rifamycins, functioning as an RNAP

protection protein. We directly tested this hypothesis *in vitro* and found that HelR<sub>SV</sub> does rescue transcription in inhibitory concentrations of rifamycins but by a comparatively small amount (2-3 fold). This does not fully recapitulate the fold change in MIC this gene confers (8-16 fold). We note that experiments with a HelR and RNAP from *Mycobacterium abscessus* show a comparable result in the same assay (P. Ghosh, personal communication).

For HelR<sub>SV</sub> to confer resistance to rifamycins by displacing them from RNAP, RNAP must bind promoter DNA and initiate transcription at a rate that exceeds rifamycin rebinding to the enzyme. In the context of a simple *in vitro* system, one would expect the rebinding rate to be very fast; this may limit the protective effect of HelR<sub>SV</sub> *in vitro*. *In vivo*, displaced rifamycins may interact weakly with other cytoplasmic proteins, likely slowing their rebinding. The concentration of rifamycins in the cytosol is further modulated by active efflux, which would also restrict the effective rate of RNAP rebinding for free rifamycins relative to our *in vitro* system. Such factors may contribute to the differences we observe for HelR<sub>SV</sub> mediated resistance *in vitro* and *in vivo*. Additionally, we cannot rule out the requirement for other protein partners, which may be required for maximal HelR<sub>SV</sub> activity. HelD<sub>BS</sub>, for instance, acts synergistically with the Firmicute specific  $\delta$  subunit of RNAP (Pei et al., 2020; Wiedermannová et al., 2014). Furthermore, rifamycin-bound RNAP also engages in abortive transcription, occupying the template DNA's promoter. These complexes can also prevent free-RNAP from binding, thereby masking the effects of HelR<sub>SV</sub> on free-RNAP *in vitro*. Nevertheless, both the in-cell and *in vitro* experiments align with HelR's importance in rifamycin resistance. It is also relevant to note

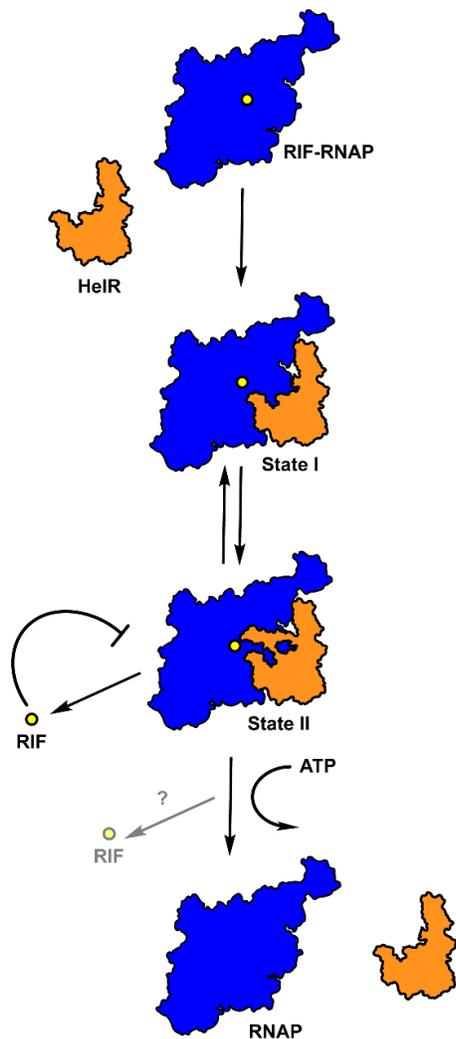
that we do not observe any stimulation of transcription by HelR<sub>Sv</sub>, offering further evidence that it is not functionally equivalent to HelD and instead has a dedicated function in antibiotic resistance.

We specifically tested the hypothesis that HelR<sub>Sv</sub> can displace rifamycins from their binding site on RNAP by designing and synthesizing a photo-crosslinking rifamycin (RPP). Structural data from HelR<sub>Ms</sub> implies that HelR and DNA cannot co-exist within RNAP, suggesting that HelR can only remove rifamycins from free-RNAP (Kouba et al., 2020; Newing et al., 2020; Pei et al., 2020). The addition of HelR<sub>Sv</sub> suppressed labeling of free RNAP in a dose-dependent manner, indicative of displacement from RNAP. Furthermore, a combination of HelR<sub>Sv</sub> and the rifamycin inactivating enzyme Arr can generate rifampin-free RNAP in solutions containing saturating rifampin concentrations as measured by their ability to transcribe when template DNA and NTPs are added. The addition of Arr alone is insufficient to rescue transcription because it cannot use RNAP bound rifampin as a substrate, by displacing rifamycins from RNAP HelR<sub>Sv</sub> allows them to be inactivated by Arr. Together these experiments provide strong evidence for the active displacement of rifamycins from RNAP by HelR<sub>Sv</sub> and provide the molecular logic for rifamycin resistance conferred by HelR.

Using RPP, we also show that ATP is not required for the displacement of rifamycins. The addition of ATP decreases labeling (increases rifamycin binding). This was surprising given that an ATPase null mutant of HelR<sub>Sv</sub> cannot confer resistance. We initially considered that ATPase activity might actively drive displacement of rifamycins. However, hydrolysis of ATP/GTP does not result in drug displacement in several classes of ribosomal

protection proteins. TetO and TetM are GTPases that bind ribosomes already complexed with GTP and use the energy released by GTP hydrolysis to dissociate from the ribosome, not to displace tetracyclines directly. Their ability to displace tetracyclines is due to conformational changes they induce in the ribosome (Dönhöfer et al., 2012; Li et al., 2013; Wilson et al., 2020). The same cycle underlies the activity of ABC-F protection proteins such as VmlR or MsrE, which use ATP instead of GTP and confer resistance to different classes of ribosome targeting antibiotics (Crowe-McAuliffe et al., 2018; Su et al., 2018). A separate cycle has been elucidated for HelD from *B. subtilis* and is likely conserved across all HelD-like proteins. Pei *et al.* observe the dissociation of RNAP-HelD<sub>Bs</sub> complexes following the addition of ATP and non-hydrolyzable analogs (such as AMPPNP/ATP $\gamma$ S), which suggests that ATP binding drives dissociation from RNAP and that the hydrolysis of ATP regenerates free HelD. Applying this model to HelR<sub>Sv</sub>, ATP would lead to the displacement of HelR from RNAP, consistent with our *in-vitro* data showing that 1mM ATP increases labeling of RNAP by RPP. We speculate that our ATPase null mutant is impaired in RNAP binding because it cannot hydrolyze bound ATP and is locked in a conformation incompatible with RNAP binding. The precise steps in the catalytic cycle of HelR and the role of ATP binding and hydrolysis require further study. Still, this work firmly establishes that HelR<sub>Sv</sub> removes rifamycins from RNAP in a manner analogous to the well-characterized ribosomal protection proteins.

These data support a molecular model of HelR-mediated resistance (**Figure 6**). Following exposure to rifamycin antibiotics, HelR is produced and forms a complex with RNAP. The structures of HelR<sub>Ms</sub> suggest that this begins with HelR in state I, with a



**Figure 6 Model of HelR-mediated rifamycin resistance.** HelR (orange) binds to free RNAP (blue) with rifamycin bound (RIF, yellow circle). HelR binds in State I and isomerizes to State II, which we hypothesize dislodges RIF. In this conformation, re-binding may also be disfavored. As in *B. subtilis* Held, ATP binding and hydrolysis likely drive the dissociation of HelR and RNAP regenerating free species capable of transcription.

disordered PCh loop, which subsequently folds into the primary channel generating state II. We hypothesize that this state disfavors rifamycin binding, resulting in their removal from RNAP. ATP hydrolysis does not drive the dissociation of rifamycins and is instead

likely required for dissociation of HelR from RNAP so transcription can occur. Rifamycins cannot inhibit RNAP once transcription has begun; by constantly removing these antibiotics from free-RNAP, HelR maintains a pool of transcriptionally-competent RNAP, thereby conferring resistance. In support of this mechanism, modeling by Kouba *et al.* show that HelR<sub>Ms</sub> can form a complex with RNAP with both  $\sigma^A$  and RpbA bound, facilitating transcription initiation following removal of rifamycins (Kouba et al., 2020).

Microbes exhibit tolerance to antibiotics when they survive a transient exposure to inhibitory levels of a bactericidal antibiotic (Brauner et al., 2016). Rifamycins show bactericidal activity against Gram-positive organisms such as the mycobacteria (Floss and Yu, 2005), making this drug class highly important for treating tuberculosis. Despite this, the mechanism by which rifamycins kill bacteria is not well understood. It has been tied to aberrant central metabolism and reactive oxygen species generation in multiple studies (Lobritz et al., 2015; Piccaro et al., 2014). Several known tolerance mechanisms have been characterized, including rifampin efflux pumps, overproduction of RNA polymerase, and even mistranslation of RpoB, leading to the generation of a subpopulation of rifamycin-resistant RNAP (Adams et al., 2011; Javid et al., 2014; Zhu et al., 2018). In this study, we demonstrate that deletion of *helR* decreases tolerance of *S. venezuelae* to rifampin. An effect that is not observed for the rifamycin inactivating enzyme Rox. HelR's ability to displace rifamycins from RNAP likely sustains some level of transcription in the presence of high concentrations of rifamycins. For ribosome targeting antibiotics, the relative degree of 'cidality' is positively correlated with slow dissociation rates from the ribosome (Svetlov et al., 2017). This parallels our findings with HelR, which dissociates rifamycins from

RNAP and accordingly decreases their bactericidal effects (increases tolerance). We note that rifamycin-mediated cell death is rapid and that many bacterial strains with a RAE-associated *helR* also have RAE-associated genes encoding rifamycin inactivating enzymes. We speculate that HelR's principal role may be to provide rifamycin tolerance following drug exposure to enable sufficient destruction of the antibiotic in the local environment by RAE-associated inactivation enzymes. In support of this hypothesis, we show that the combination of Arr and HelR<sub>SV</sub> can rescue transcription in high concentrations of rifamycins *in-vitro* (**Figure 4G**). This would explain why many bacterial strains have both HelR and rifamycin inactivating enzymes under control of RAEs – HelR's provide tolerance in the presence of the antibiotic, offering sufficient time for inactivating enzymes to decrease the concentration of drugs below the MIC.

HelD-like enzymes are numerous in the Firmicute and Actinobacteria phyla. We show that a small fraction of these are associated with RAEs and, therefore, likely contribute to rifamycin resistance. Much of the diversity within this large family of proteins has yet to be interrogated. Some may be involved in resistance to other RNAP inhibitors, or they may function to aid in transcriptional cycling like HelD. Of the fraction engaged in resistance, many copies are found in pathogenic bacteria. A particularly relevant example is *M. abscessus* (P. Ghosh, personal communication), but HelRs are present in many fast-growing mycobacteria. These genes are also found in pathogenic *Nocardia* species such as *N. farcinia* and the foal pathogen *Rhodococcus equi*. The identification and understanding of HelR can now guide the synthesis of rifamycin analogs that are not

substrates for HelR, leading to more effective rifamycin chemotherapy for various infections in the resistance era.

### **Limitations of the study**

This work establishes that HelR confers rifamycin resistance by causing the dissociation of rifamycins from free-RNAP, but several questions remain. As mentioned previously, we could not determine the structural basis of rifamycin displacement from the co-structure(s) of HelR<sub>MS</sub>-RNAP, and the precise role of NTP binding and hydrolysis in HelR-mediated resistance requires further investigation. One aspect of HelR function we did not consider in this work is its ability to remove rifamycin-RNAP-DNA complexes, as HelD-like enzymes are known to displace stalled RNAP from DNA. Although we do not believe that the ability to remove rifamycin-RNAP-DNA complexes is required to explain the resistance conferred by HelR, it may be another avenue by which HelR protects RNAP. We also performed our *in-vitro* transcription and labelling reactions using what amounts to the minimal essential system. It remains unclear how HelR may interact with other transcription factors,  $\sigma$ -factors, and different promoters *in-vivo*.

### **Acknowledgments**

We thank P. Ghosh for sharing preliminary data on HelR from *M. abscessus*. This work was funded by the Canadian Institutes of Health Research (FRN-148463) and by a Canada Research Chair to G.D.W. M.D.S. is supported by a Graduate Scholarship from the Natural Sciences and Engineering Council of Canada.

### **Author contributions**

M.D.S conceived and performed experiments and wrote the manuscript. K.K and N.W. performed experiments and provided expertise and feedback. G.D.W conceived of experiments, wrote, and edited the manuscript.

### **Declaration of Interests**

The authors declare no competing interests.

### **STAR METHODS**

#### **Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gerard Wright ([wrightge@mcmaster.ca](mailto:wrightge@mcmaster.ca))

#### **Materials availability**

All strains and recombinant DNA generated in this study are available upon request. Rifamycin photoprobe is available upon reasonable request. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

#### **Data and code availability**

- Protein abundance data used to generate **Figure 3A** are supplied as **Supplementary Item 1\***. FASTA files containing all proteins used to prepare **Figure 5** can be found in **Supplementary Item 2\***. Raw image data for all blots and gels have been deposited in Mendeley Data. All are publicly available at the time of publication with DOIs listed in the key resources table.
- All original code used to generate the HeID-like protein SSN are publicly available at Zenodo, the DOI is listed in the key resource table and is publically available at the time of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**\*Note that Supplementary Items 1 and 2 could not be reproduced here due to size constraints but are available in the online version of this manuscript.**

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
THE™ DYKDDDDK Tag Antibody [HRP], mAb, Mouse	Genscript	Cat# A01428; Clone ID 5A8E5
RNA polymerase beta Monoclonal Antibody	Invitrogen	Cat# MA1-25425; Clone ID 8RB13
Rabbit Anti-Mouse IgG H&L (HRP)	Abcam	Cat# ab97046
Streptavidin-HRP Conjugate	Cytiva	Cat# GERPN1231-2ML
<b>Bacterial and virus strains</b>		
<i>Streptomyces venezuelae</i>	ATCC	ATCC10712
$\Delta$ helR	This paper	N/A
$\Delta$ rox	This paper	N/A
$\Delta$ helR-scar	This paper	N/A
$\Delta$ helR pIJ10257	This paper	N/A
$\Delta$ helR pIJ:helR	This paper	N/A
$\Delta$ helR $\Delta$ rox	This paper	N/A
$\Delta$ helR $\Delta$ rox pIJ10257	This paper	N/A
$\Delta$ helR $\Delta$ rox pIJ:helR	This paper	N/A
$\Delta$ helR $\Delta$ rox pIJ:helR <sub>D533A</sub>	This paper	N/A
$\Delta$ helR $\Delta$ rox pIJ:helR-his <sub>6</sub>	This paper	N/A
$\Delta$ helR $\Delta$ rox pIJ:helR-FLAG	This paper	N/A
pGUS:P <sub>rox</sub>	This paper	N/A
$\Delta$ helR pGUS:P <sub>rox</sub>	This paper	N/A
$\Delta$ rox pGUS:P <sub>rox</sub>	This paper	N/A
$\Delta$ helR $\Delta$ rox pGUS:P <sub>rox</sub>	This paper	N/A
<b>Bacillus subtilis 168</b>	ATCC	ATCC23857
RpoB H482Y	(Spanogiannopoulos, 2014)	N/A
$\Delta$ helD	(Koo et al., 2017)	Bacillus Genetic Stock Center: BKK33450
<i>Amycolatopsis mediterranei</i> S699	Dr. Tin-Wein Yu (August et al., 1998)	N/A
<i>Escherichia coli</i> DH5 $\alpha$	ThermoFisher	Cat#18265017
<i>Escherichia coli</i> BL21 (DE3) pLysS	Novagen	Cat: 69451-3
pET28a:hrdB	This paper	N/A
<i>Escherichia coli</i> ET12567	Gift from Dr. MJ Buttner (MacNeil et al., 1992)	N/A
pUZ8002::bla	(Cherepanov and Wackernagel, 1995)	N/A
<i>Escherichia coli</i> DH5 $\alpha$ pCP20	(Cherepanov and Wackernagel, 1995)	N/A
<i>Escherichia coli</i> BW25113 pKD46	(Datsenko and Wanner, 2000)	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
Rifampin	Sigma Aldrich	Cat# R3501; CAS: 13292-46-1
Rifaximin	Sigma Aldrich	Cat# 33999; CAS: 80621-81-4
Rifabutin	Sigma Aldrich	Cat# R3530; CAS: 72559-06-9
Rifamycin SV	Sigma Aldrich	Cat# PHR1664; CAS: 14897-39-3
Rifamycin B	This paper	CAS 13929-35-6
Fidaxomicin	Sigma Aldrich	Cat# SML1750; CAS Number: 873857-62-6

Vancomycin	Sigma Aldrich	Cat# V2002; CAS Number: 1404-93-9
Tetracycline	Sigma Aldrich	Cat# 87128; CAS Number: 60-54-8
X-glucuronide	Alfa Aesar	Cat# J64360-03; CAS Number 129541-41-9
ANTI-FLAG® M1 Agarose Affinity Gel	Sigma Aldrich	Cat# A4596
FLAG peptide	Genscript	Cat# RP10586
5-hexynoic acid	Sigma Aldrich	Cat# 544000; CAS Number: 53293-00-8
4,4'-diaminobenzophenone	Sigma Aldrich	Cat# 8.15111; CAS Number: 611-98-3
1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDAC)	Sigma Aldrich	Cat# 341006; CAS Number: 25952-53-8
Hydroxybenzotriazole (HOBt)	Sigma Aldrich	Cat# 54802; CAS Number: 123333-53-9
Azide-PEG3-biotin	Sigma Aldrich	Cat# 762024; CAS Number:875770-34-6
Rifamycin photoprobe (RPP)	This paper	N/A
dms0-d6-100%	Cambridge Isotopes	Cat# DLM-34 10x0.75
NTP set, 100mM solution	Thermo Scientific	Cat# R0481
20 µCi[α- <sup>32</sup> P] UTP (800 Ci/mmol)	Perkin Elmer	Cat# BLU507T250UC
<b>Critical commercial assays</b>		
SYBR Select Master Mix CFX	Applied Biosystems	Cat# 4472937
Maxima First Strand cDNA synthesis kit	Thermo Scientific	Cat# K1671
Pierce™ BCA Protein Assay Kit	Pierce	Cat# 23225
SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Scientific	Cat# 34580
<b>Deposited data</b>		
Raw image data, blots and gels	Mendeley Data	DOI: 10.17632/rjed7phswc.1
Original code used to generate HelD-like protein SSN	Zenodo	DOI: 10.5281/zenodo.6573906
<b>Oligonucleotides</b>		
See <b>Table S4</b>	All Oligos ordered from IDT	N/A
<b>Recombinant DNA</b>		
pIJ10257	Dr. MJ Buttner, (Hong et al., 2005)	N/A
pIJ: <i>helR<sub>sv</sub></i>	This paper	N/A
pIJ: <i>helR<sub>sv</sub>-his<sub>6</sub></i>	This paper	N/A
pIJ: <i>helR<sub>sv</sub>-FLAG</i>	This paper	N/A
pIJ: <i>helR<sub>sv</sub> D533A</i>	This paper	N/A
pGUS	Dr. MA Elliot (Myronovskyi et al., 2011)	N/A
pGUS: <i>P<sub>rox</sub></i>	This paper	N/A
SV-2 G10	Dr. MJ Buttner-cosmids available at <a href="http://strepdb.streptomyces.org.uk/">http://strepdb.streptomyces.org.uk/</a>	N/A
SV-2 C08	Dr. MJ Buttner	N/A
SV-5 E02	Dr. MJ Buttner	N/A
SV-2 G10Δ <i>rox</i>	This paper	N/A
SV-2 C08Δ <i>helR</i>	This paper	N/A

SV-5 E02 $\Delta$ SVEN5092	This paper	N/A
SV-2 C08 $\Delta$ helR-scar	This paper	N/A
pET28a	EMD biosciences	Cat# 69864-3
pET28a:hrdB	This paper	N/A
pCAP03-aac3(IV)	Dr. Bradley Moore (Tang et al., 2015)	RRID: Addgene_69862
pIJ10701	Dr. MJ Buttner (Gust et al., 2003)	N/A
P <sub>AP3</sub> gBlock (Figure S4)	IDT	N/A
<b>Software and algorithms</b>		
ImageJ v1.53g	NIH	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>
Snakemake v6.0.5	(Köster and Rahmann, 2018)	<a href="https://snakemake.readthedocs.io/en/v6.0.5/">https://snakemake.readthedocs.io/en/v6.0.5/</a>
NCBI genome download v0.3.0	Dr. Kai Blin	<a href="https://github.com/kblin/ncbi-genome-download">https://github.com/kblin/ncbi-genome-download</a>
HMMER 3.1b	(Eddy, 2011)	<a href="http://hmmmer.org/download.html">http://hmmmer.org/download.html</a>
Usearch v11.0.667	(Edgar, 2010)	<a href="https://drive5.com/usearch/download.html">https://drive5.com/usearch/download.html</a>
Diamond v0.9.14.115	(Buchfink et al., 2014)	<a href="https://github.com/bbuchfink/diamond">https://github.com/bbuchfink/diamond</a>
Mafft v7.310	(Katoh and Standley, 2013)	<a href="https://mafft.cbrc.jp/alignment/software/">https://mafft.cbrc.jp/alignment/software/</a>
TrimA1 v1.4.rev22	(Capella-Gutiérrez et al., 2009)	<a href="http://trimal.cgenomics.org/">http://trimal.cgenomics.org/</a>
Cytoscape v3.8.2	(Shannon et al., 2003)	<a href="https://cytoscape.org/">https://cytoscape.org/</a>
CDsearch	(Lu et al., 2020)	<a href="https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi">https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi</a>
Prism v7.05	Graphpad	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
Adobe Illustrator CC	Adobe	<a href="https://www.adobe.com/ca/products/illustrator.html">https://www.adobe.com/ca/products/illustrator.html</a>
ChemDraw Professional 18.0	PerkinElmer	<a href="https://www.perkinelmer.com/category/chemdraw/">https://www.perkinelmer.com/category/chemdraw/</a>
Peaks Studio X	Bioinformatics Solutions	<a href="https://www.bioinform.com/">https://www.bioinform.com/</a>
Scaffold 5	Proteome software	<a href="https://www.proteomesoftware.com/products/scaffold-5/">https://www.proteomesoftware.com/products/scaffold-5/</a>
<b>Other</b>		
HiPrep Heparin FF 16/10 column	Cytiva	Cat# 28936549
Superdex 200 10/300 GL column	Cytiva	Cat# 17517501
Capto HiRes Q 10/100 column	Cytiva	Cat# 29275881
RediSep <sup>®</sup> RF Silver C18 26g column	Teledyne	Cat# 69-2203-412
RediSep <sup>®</sup> RF Silver C18 4.3g column	Teledyne	Cat# 69-2203-410

## Experimental Model and Subject details

*Streptomyces venezuelae* ATCC 10712 was grown using Bennett's media (10 g potato starch, 2 g casamino acids, 1.8 g yeast extract, and 2 mL of Czapek's mineral mix [10 g KCl, 10 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 12 g NaNO<sub>3</sub>, 0.2 g FeSO<sub>4</sub>, 200 µL concentrated HCl in 100 mL H<sub>2</sub>O] per 1 L H<sub>2</sub>O) or Tryptic Soy Broth (BD Bacto™) with and without 1.5% agar

depending on the application. Biparental matings took place on Soy Flour Mannitol (20 g soy flour, 20 g mannitol, and 20 g agar) plates supplemented with 15 mM MgCl<sub>2</sub>. *Escherichia coli* was grown using LB (Lysogeny broth). *Bacillus subtilis* was grown using Tryptic Soy Broth (BD Bacto™). Antibiotics for selection were used at the following concentrations 7.5 µg/mL kanamycin, 100 µg/mL ampicillin, 34 µg/mL chloramphenicol, 50 µg/mL apramycin, 100 µg/mL and 50 µg/mL hygromycin B for *E. coli* and *S. venezuelae*, respectively. *S. venezuelae* and *B. subtilis* were grown at 30°C and *E. coli* was grown at 37°C.

### **RNA Isolation and cDNA synthesis**

Approximately 10<sup>9</sup> spores of *S. venezuelae* were used to inoculate a 50 mL starter culture in Bennett's media. Following incubation for 16h at 30°C with shaking (250 rpm), the culture was used to seed 6 x 200 mL subcultures (with a 1:100 dilution of starter). These subcultures represent two sets of biological triplicates, one receiving rifamycin SV treatment (0.5 µg/mL) and vehicle control (DMSO). After 8 hours of growth, the cultures were treated and returned to the incubator for another two hours, after which 10 mL of culture were collected in a 15 mL Falcon tube and centrifuged for 10 minutes at 8 000 x g at 4°C. The supernatant was discarded, and the pellets were flash-frozen in liquid nitrogen and stored at -80°C until RNA isolation. Cell pellets were incubated with 300 µL of 10 mg/mL lysozyme (prepared fresh in sterile H<sub>2</sub>O) and incubated for 7 minutes at 30°C/250 rpm to allow for efficient lysis by TRIzol reagent (Ambion). 10-15 glass beads (4mm diameter) and 4 mL of TRIzol reagent were added to each cell pellet. Samples were then vortexed for 5 minutes, and 0.8 mL of chloroform was added, followed by 4 cycles of 30

seconds vortexing followed by 30 seconds on ice. The glass beads were removed, and the phenol/chloroform mixture was separated using centrifugation (10 minutes at 8 000 x g and 4°C). The aqueous phase was aliquoted to a fresh falcon tube and an equal volume of phenol-chloroform (pH 4, Ambion) was added. This solution was vortexed for 30s and centrifuged once more. The aqueous phase was mixed 1:1 with 75% ethanol and applied to a column from the Purelink™ RNA Mini Kit (Invitrogen). The rest of the purification was performed according to the manufacturer's instructions. RNA integrity was qualitatively examined by gel electrophoresis. cDNA was synthesized from 1 µg of RNA using the Maxima First Strand cDNA synthesis kit (Thermo Scientific) according to the protocol for high GC content templates.

### **RT-qPCR**

Quantitative PCR was performed with SYBR Select Master Mix CFX (Applied Sciences) on a Bio-Rad C1000 thermocycler. All primer concentrations were 200 nM, cDNA was diluted 1:1 with nuclease-free water (Ambion), and 2 µL of diluted cDNA was used as the template. Thermocycling conditions were as follows, 2 minutes at 50 °C, 2 minutes at 95 ° C, then 40 cycles of (95 °C for 15s, 57 °C for 15s, and 72 °C for 30s). Quantitative PCR was performed on a Bio-Rad C1000 Thermocycler. Gene-specific primers are as follows *hrdB*-q FP/RP, *rox*-q FP/RP, *helR*-q FP/RP. Expression is reported as a ratio of the housekeeping gene *hrdB*. P-values were determined using an unpaired students t-test from 3 independent biological replicates for induced and non-induced cultures.

## Mutant construction

Deletions of *rox* (*SVEN\_0481*), *helR* (*SVEN\_6029*), and *SVEN5092* were made using the REDIRECT PCR targeting method (Gust et al., 2003). *E. coli* BW25113 pKD46 was first transformed with cosmids SV-2 G10 (33-45kb insert of *S. venezuelae* gDNA containing *rox*), SV-2 C08 (cosmid containing *helR*), and SV-5 E02 (cosmid containing *SVEN5092*). Cosmids used in this work are from a genomic library of *S. venezuelae* which can be accessed at <http://strepdb.streptomyces.org.uk>. Using primer pairs *rox* K.O. FP/RP, *helR* K.O. FP/RP, and *SVEN5092* K.O. FP/RP we amplified an *oriT*-apramycin resistance cassette flanked by FRT sites with 39bp of homology to the 5' and 3' ends of the genes to be deleted. In frame deletions were generated by  $\lambda$  red recombination with these PCR amplicons and their respective cosmids. The *oriT* apramycin resistance cassette was amplified using pCAP03-aac3(IV) as template (Tang et al., 2015). Recombinant cosmids were confirmed by PCR using primer pairs flanking the target gene (*rox* cPCR FP/RP, *helR* cPCR FP/RP, *SVEN5092* cPCR FP/RP). Cosmids containing disrupted genes (SV-2 G10 $\Delta$ *rox*, SV-2 C08 $\Delta$ *helR*, and SV-5 E02 $\Delta$ *SVEN5092*) were transformed into *E. coli* ET12576 pUZ8002::*bla* and introduced to *S. venezuelae* by biparental mating (Kieser et al., 2000). Successful recombination was identified by ex-conjugants with an Apra<sup>r</sup>/Kan<sup>s</sup> sensitivity profile and confirmed using *rox* cPCR FP/RP, *helR* cPCR FP/RP, *SVEN5092* cPCR FP/RP. For *rox* and *helR* double recombinants were obtained immediately following conjugation, but *SVEN5092* yielded only Apra<sup>r</sup>/Kan<sup>f</sup> ex-conjugants (single recombinants) and had to be passaged on non-selective media until a double recombinant could be isolated. To construct *S. venezuelae*  $\Delta$ *helR* $\Delta$ *rox* we removed the *oriT* apramycin cassette

from mutagenized SV-2 C08 by introducing this cosmid into *E. coli* DH5 $\alpha$  pCP20, which expresses FLP recombinase. Removal of this cassette was confirmed by PCR using *helR* cPCR FP/RP, this verified cosmid was then re-introduced into *E. coli* BW25113 pKD46 and an *oriT* hygromycin B resistance cassette was then inserted into the cosmid backbone using PCR targeting. The *oriT* hygromycin cassette was amplified from pIJ10701 using the primers *bla* FP/RP. This cosmid was introduced into *S. venezuelae*  $\Delta helR$  by conjugation using *E. coli* ET12567 pUZ8002::*bla* and selecting for single recombination (Hyg<sup>r</sup>). Individual colonies were screened until a double recombinant (Apra<sup>s</sup>/Hyg<sup>s</sup>) was found; we termed this strain *S. venezuelae*  $\Delta helR$ -scar and confirmed the excision of the *oriT* apramycin cassette by PCR. We could now use mutagenized SV-2 G10 as described above to generate *S. venezuelae*  $\Delta helR\Delta rox$  because *S. venezuelae*  $\Delta helR$ -scar is Apra<sup>s</sup>. *S. venezuelae*  $\Delta helR$ -scar was also used to construct *S. venezuelae*  $\Delta helR\Delta 5092$  in the same manner.

### **Complementation and HelR cloning**

In order to complement *S. venezuelae* *helR*<sup>-</sup> strains described in this work, we amplified *helR* from *S. venezuelae* genomic DNA using the primers *helR* FP/RP. The resulting amplicon was digested using NdeI and HindIII and ligated into pIJ10257, an *E. coli* – *Streptomyces* integrative shuttle vector with a constitutive *Streptomyces* promoter P<sub>ermE</sub>\* and RBS directly upstream of the NdeI site (Hong et al., 2005). Tagged *helR* constructs were generated by PCR using *helR* FP and a long reverse primer which encodes for a C-terminal his<sub>6</sub> tag (*helR*-his<sub>6</sub> RP) or a C-terminal FLAG tag (*helR*-FLAG RP) and a new stop codon, these amplicons were then digested with NdeI and HindIII and ligated into

pIJ10257. The D533A substitution in HelR<sub>SV</sub> was generated by overlap extension PCR where *helR* was amplified from pIJ10257:helR<sub>SV</sub> in two fragments by the primers *helR* D533A FP1/RP1 and FP2/RP2 respectively (FP1 and RP2 bind to plasmid sequence). These products overlap at RP1 and FP2 where the 533Asp codon GAC is replaced with GCC for Ala by including this change in both primers. 15 cycles of PCR was performed with the amplicons from FP1/RP1 and FP2/RP2, joining these two fragments, and an additional 20 cycles was performed after the addition of FP1 and RP2 to generate full length product containing *helR* D533A. This amplicon was digested using NdeI and HindIII and ligated into pIJ10257. A colorimetric reporter for RAE induction was created using the pGUS vector which contains a promoter-less  $\beta$ -glucuronidase gene (*GUS*), the intergenic region upstream of *rox* was amplified from *S. venezuelae* genomic DNA using the primers P<sub>rox</sub> FP/RP, the product was digested using XbaI and KpnI and ligated into pGUS, directly upstream of *GUS* (Myronovskyi et al., 2011). All constructs were confirmed by sanger sequencing, pIJseq FP/RP and HelRint were used for pIJ10257:*helR* and related constructs and pGUSseq was used to confirm pGUS:P<sub>rox</sub>. Sequence verified constructs were transformed into *E. coli* ET12567 pUZ8002::*bla* and introduced to *S. venezuelae* strains by biparental mating.

### **Antibiotic susceptibility testing**

Antimicrobial susceptibility testing was performed according to general broth microdilution CLSI protocols in 96 well plates (Weinstein et al., 2018). *S. venezuelae* inoculum for MICs was grown in 3mL TSB with 2-3 4mm glass beads for 48 hours at 30°C with shaking (250rpm). Under these conditions, *S. venezuelae* does not form large

clumps/aggregates, and cultures can be reliably standardized by OD<sub>600nm</sub>. All *S. venezuelae* MICs were determined in TSB following incubation at 30°C for three days with shaking (250rpm), *B. subtilis* MICs were determined after 24h of incubation under the same conditions. *B. subtilis*  $\Delta$ hcd was obtained from the ordered knockout collection and is the kanamycin resistant clone (BKK33450) (Koo et al., 2017). All antibiotics (Rifampin, Rifaximin, Rifabutin, Rifamycin SV, Fidaxomicin, Vancomycin, and Tetracycline) were all purchased from Sigma Aldrich.

### **GUS assays and agar-based susceptibility testing**

*S. venezuelae* strains were grown as described for antibiotic susceptibility testing and then standardized to an OD<sub>600nm</sub> of 0.1 and a sterile cotton swab was used to inoculate a lawn onto a plate of Bennett's media containing 80 µg/mL X-glucuronide (Alfa Aesar). Sterile cellulose discs containing, 1 µg Rifampin, 50 µg Kanamycin, or 10 µL DMSO as a control were added on top of the lawn of *S. venezuelae* and the plates were incubated at 30°C for 48 hours before imaging. To characterize the bioactivity of RPP cellulose discs containing 20 µg or 40 µg RPP were placed onto lawns of *B. subtilis* and *S. venezuelae* strains respectively, prepared in the manner described above. *S. venezuelae* was inoculated onto Bennett's media and *B. subtilis* was inoculated onto TSA and both incubated at 30 °C. *S. venezuelae* plates were imaged after 48 h incubation and *B. subtilis* after 24h. *B. subtilis* H482Y was generated previously (Spanogiannopoulos, 2014).

### **Antibiotic time-kill assay**

Spores of *S. venezuelae* were cultured overnight in 3mL TSB, then diluted 1:100 into a fresh flask of 50mL of TSB and allowed to grow for 6 hours to reach exponential

phase. These cultures were standardized to an OD<sub>600nm</sub> of 0.1, and rifampin (10X MIC of each strain) or DMSO was added. At each specified time point (0h, 8h, and 24h), 1mL of culture was harvested by centrifugation (10 000 x g 3 minutes) and resuspended in 1mL of fresh TSB. Centrifugation and resuspension were repeated two additional times to remove all residual rifampin. Serial 10-fold dilutions of culture were then spotted onto Bennett's agar and incubated for 24 hours before imaging. Bennett's agar was chosen over TSA for this application because *S. venezuelae* forms large amorphous colonies on TSA making spot dilutions challenging to interpret.

### ***S. venezuelae* RNA polymerase purification**

The same growth conditions established for RT-qPCR were used to grow *S. venezuelae* for RNAP isolation under inducing/non-inducing conditions, these preparations were analyzed **Figure 3A**. RNAP used for *in-vitro* studies was purified from *S. venezuelae*  $\Delta$ *helR* to prevent any possible contamination by HelR<sub>Sv</sub> and TSB was used for cultivation in place of Bennett's media because it yielded more biomass per liter.

Native RNA polymerase was purified from *S. venezuelae* using the procedure outlined in (Kieser et al., 2000) with minor modifications, based on the protocol from Burgess and Jendrisak (Burgess and Jendrisak, 1975). Mycelia were collected from liquid culture by centrifugation at 8000 x g for 15 minutes and ~50g pellets were stored at -80°C until purification. A single 50g pellet was resuspended in 200mL of Grinding buffer (50mM Tris pH 7.9, 5% (v/v) glycerol, 2mM EDTA, 0.233M NaCl, 1mM  $\beta$ -mercaptoethanol, 0.1mM DTT, 23  $\mu$ g/mL PMSF) and lysed by two consecutive passages through a continuous flow cell disruptor (Constant Systems Limited, Daventry U.K.) at

30K psi. The second passage is necessary to completely homogenize the lysate and reduce the viscosity. Next, 266mL of TGED (10mM Tris pH7.9, 5% glycerol (v/v), 0.1mM EDTA, 0.1mM DTT) + 0.2M NaCl is added to the lysate and fresh PMSF is added to maintain the concentration at 23 µg/mL. This solution was centrifuged for 15 minutes at 20 000 x g at 4°C. The supernatant was collected, and the volume was measured. Avoiding any carryover of the pellet at this step is important and occasionally necessitated a second abbreviated centrifugation step (5 minutes) to recover all the supernatant. 10% (w/v) Polyethylenimine (PEI) was added drop by drop to the supernatant with constant stirring, this step was performed at 4°C and 3.5mL of 10% PEI were added per 100mL of supernatant. Once all the PEI has been added the solution was left to continue mixing for another 10 minutes before collecting insoluble material by centrifugation at 4000 x g for 90s. At this NaCl concentration, RNAP is found in the pellet, so the supernatant can be safely discarded. The pellets were resuspended in 400mL TGED + 0.5M NaCl using a 10mL serological pipette and were left to stir at 4°C for 15 minutes before centrifuging once more at 4000 x g for 90s. The supernatant was discarded, and the pellets were resuspended in 200mL TGED + 1M NaCl and allowed to stir for 20 minutes at 4°C. At this NaCl concentration RNAP will enter solution, insoluble material is removed by centrifugation at 20 000 x g for 10 minutes and the supernatant is collected. Protein was then precipitated by the addition of 35g of ammonium sulfate for each 100mL of supernatant, ammonium sulfate was added 1-2g at a time with constant stirring at 4°C. After 20 minutes of additional stirring after all the ammonium sulfate has been added the precipitate was collected by centrifugation at 35 000 x g for 30 minutes at 4°C. Pellets were resuspended in TGED10 (10mM Tris pH7.9, 10%

glycerol (v/v), 0.1mM EDTA, 0.1mM DTT) which was added until the solution was brought to equal conductivity with TGED10 + 0.15M NaCl, usually ~100mL. 50mL of this solution was loaded at a flow rate of 2 mL/min onto a HiPrep Heparin FF 16/10 column (Cytiva), using an AKTA pure FPLC system. The column was washed with 20 column volumes (CVs) of TGED10 + 0.15M NaCl at a flow rate of 10 mL/min and protein was step-eluted with 5CVs of TGED10 + 1M NaCl. 2-3 consecutive runs were usually necessary to process the entire sample. All protein containing elution fractions (as judged by  $A_{280\text{nm}}$ ) were pooled and precipitated once more using ammonium sulfate as described above. Precipitate was harvested by centrifugation at 35 000\_x g for 30 minutes at 4°C and protein pellet was resuspended in 0.8-1mL TGED + 0.15M NaCl. Multiple 250  $\mu$ L injections of this protein were fractionated on a Superdex 200 10/300 GL column (Cytiva) equilibrated in TGED + 0.15M NaCl with a flow rate of 0.25mL/min. Fractions containing RNAP (as determined by SDS-PAGE) were pooled. RNAP prepared for proteomic analysis was dialyzed into 50mM  $\text{NaH}_2\text{PO}_4$ , pH7.0, 0.1mM EDTA, 0.1mM DTT, 0.5M NaCl for compatibility with downstream procedures and then concentrated using 70kDa cut-off Amicon Ultracentrifugal filters (EMD Millipore). RNAP used for *in-vitro* transcription and labelling with RPP was further purified by anion exchange using a Capto HiRes Q 10/100 column (Cytiva). Semi-pure RNAP was loaded ~5mg at a time with a flow rate of 0.25 mL/min, and the column was washed with 10CV of TGED + 0.25M at 0.5 mL/min, and RNAP was eluted over a shallow gradient of 0.25M NaCl to 0.4M NaCl over 30 CV at 0.5 mL/min. Fractions containing pure RNAP were determine by SDS-PAGE and then pooled, concentrated, and mixed 1:1 with glycerol and stored at -20°C.

## Proteomics

100µg of semi-pure RNAP from rifampin induced and non-induced cultures were reconstituted in 50 mM ammonium bicarbonate with 10 mM TCEP [Tris(2-carboxyethyl)phosphine hydrochloride; Thermo Fisher Scientific], and vortexed for 1 h at 37°C. Chloroacetamide (Sigma-Aldrich) was added for alkylation to a final concentration of 55 mM. Samples were vortexed for another hour at 37°C. 1 µg trypsin (Pierce, MS grade) was added, and digestion was performed for 8 h at 37°C. Peptides were then dried and solubilized in 5% acetonitrile with 4% formic acid. The samples were loaded on a 1.5 µL pre-column (Optimize Technologies) and peptides were separated on a home-made reversed-phase column (150-µm internal diameter by 200 mm, packed with C18 resin) with a 56-min gradient from 10 to 30% acetonitrile (0.2% formic acid) and a 600-nl/min flow rate on an Easy nLC-1000 (Thermo Fisher Scientific) connected to a Q-Exactive HF (Thermo Fisher Scientific). Each full MS spectrum acquired at a resolution of 60,000 was followed by tandem-MS (MS-MS) spectra acquisition on the 15 most abundant multiply charged precursor ions. Tandem-MS experiments were performed using higher energy collision dissociation (HCD) at a collision energy of 27%. The data were processed using PEAKS X (Bioinformatics Solutions) and the Uniprot database corresponding to *Streptomyces venezuelae* ATCC10712 (ProteomeID: UP000006854, 7,451 proteins). Mass tolerances on precursor and fragment ions were 10 ppm and 0.01 Da, respectively. Fixed modification was carbamidomethyl (C). Variable selected posttranslational modifications were acetylation (N-ter), oxidation (M), deamidation (NQ), phosphorylation (STY). The

data were visualized with Scaffold 5.0 (Protein threshold, 99%, with at least 2 peptides identified and a false-discovery rate [FDR] of 1% for peptides). Precursor intensity quantitation was performed with PEAKS X to determine relative protein abundance across samples (Raw precursor intensity data for all 76 proteins identified in both samples is presented in **Supplemental file 1**).

### **Co-Immunoprecipitation assays**

*S. venezuelae*  $\Delta$ *helR* $\Delta$ *rox* pIJ:*helR* and pIJ:*helR*-FLAG spores were used to inoculate 3mL of TSB cultures. The next day these cultures were used to inoculate 50 mL of fresh TSB media, which was incubated for 16h before the cells were harvested by centrifugation at 10 000 x g for 10 minutes. Cells were resuspended in Tris-buffered saline (TBS) with 1 mg/mL lysozyme (Bioshop), 1 mg/mL DNase (Bovine Pancreas, Sigma Aldrich), and 1 Peirce Protease Inhibitor tablet (Thermo Fisher) and subsequently lysed by two passages through a cell disruptor (Constant Systems Limited, Daventry U.K.) at 30k PSI. Insoluble protein was removed by centrifugation at 30 000 x g for 30 minutes at 4°C. Soluble protein was quantified using a Pierce<sup>TM</sup> BCA Protein Assay Kit (ThermoFisher) and standardized to 2 mg/mL. 1mL (2 mg) of protein was added to 20  $\mu$ L of ANTI-FLAG<sup>®</sup> M1 Agarose Affinity Gel (Sigma Aldrich) pre-equilibrated in TBS and incubated on a nutator for 2 hours at room temperature. Beads were collected by centrifugation (6 000 x g for 30 seconds), the supernatant was discarded, and the beads were washed with 1mL of fresh TBS. Four additional washes of the beads were carried out as described. 60  $\mu$ L of 300  $\mu$ g/mL FLAG peptide (GenScript) was used to elute protein from the beads.

Eluted protein was separated on an 8% polyacrylamide gel and transferred onto PolyScreen PVDF transfer membrane (Perkin Elmer). Membranes were blocked overnight in TBS containing 5% Bovine Serum Albumin (Sigma Aldrich). After three 10-minute washes with TBST (0.1% Tween 20), membranes were probed with a 1:5000 dilution of an  $\alpha$ -FLAG HRP (THE™ DYKDDDDK Tag Antibody [HRP], GenScript) conjugate for detection of HelR<sub>SV</sub>-FLAG, or with a 1:5000 dilution of mouse  $\alpha$ -RpoB (8RB13, Invitrogen) followed by a 1:20000 rabbit  $\alpha$ -mouse HRP conjugate (Ab97046, Abcam) for detection of RNA polymerase. Western blots were imaged using SuperSignal™ West Pico PLUS chemiluminescence detection reagents per the manufacturer's instructions (Fisher Scientific), on a BioRad ChemiDoc with auto-exposure settings.

### **HrdB purification**

The primer set *hrdB* FP/RP was used to amplify *hrdB* from *S. venezuelae* genomic DNA and was cloned into pET28a using NdeI and HindIII. Sequence verified plasmid was transformed into *E. coli* BL21 (DE3) pLysS (Novagen). An overnight culture of *E. coli* BL21 (DE3) pLysS pET28a:*hrdB* was diluted 1:100 into 4L LB and grown at 37°C with shaking at 250rpm until an OD<sub>600nm</sub> 0.6, the culture was cooled to 16°C and induced with 0.1mM IPTG. Following a 10-hour induction, cells were collected by centrifugation 6 000 x g for 15 minutes at 4°C. Cells were resuspended into lysis buffer (25 mM HEPES pH 7.4, 400 mM NaCl, 15 mM imidazole, 1 mM DTT) with a protease inhibitor tablet, 1 mg/mL lysozyme, and 1 mg/mL DNase. Cells were lysed on a cell disruptor (Constant Systems Limited, Daventry U.K.) at 20k PSI, and soluble protein was isolated by centrifugation at 30 000 x g for 30 minutes at 4°C. Soluble protein was mixed with 2.5 mL Ni-NTA resin

(Qiagen) for 1h at 4°C on a nutator. Resin was collected by packing into a column and extensively washed with lysis buffer until protein could not be detected in the flow-through by Bradford Assay. HrdB was eluted from the column by successive washes with lysis buffer containing increasing concentrations of imidazole (50 mM, 100 mM, 200 mM, 300 mM). Fractions containing HrdB were pooled and dialyzed against lysis buffer with no imidazole. The His-tag was cleaved by overnight incubation with thrombin (from bovine plasma, Sigma Aldrich) at 4°C during dialysis. Tag-free HrdB was recovered by passage through a Ni-NTA column, concentrated to ~500uL, and purified further by injection onto Superdex 200 Increase 30/100 GL equilibrated in 25mM HEPES pH7.4, 250 mM NaCl, 1mM DTT and operated at 0.25mL/min. Fractions containing pure HrdB (determined by SDS-PAGE) were pooled and concentrated using an Amicon Centrifugal filter with 50kDa MW cutoff (EMD Millipore), snap-frozen in liquid nitrogen, and stored at -80°C.

### **HelR<sub>sv</sub> purification**

4L of TSB were inoculated with 1:500 dilution of *S. venezuelae*  $\Delta$ *helR* pIJ:*helR-his<sub>6</sub>* overnight culture and incubated at 30°C for 16 hours with shaking (250 rpm). Cells were collected by centrifugation (8 000 x g for 15 minutes at 4°C ). Lysis was carried out as described during Co-IP, except cells were resuspended in lysis buffer (25 mM HEPES pH 7.4, 400 mM NaCl, 15 mM imidazole, 1 mM DTT). Ni-NTA chromatography was performed in the same manner as for HrdB; fractions containing pure HelR<sub>sv</sub> were pooled and dialyzed against Lysis buffer lacking imidazole, then concentrated, snap-frozen, and stored at -80 °C until use.

### ***In-vitro* transcription**

Multiple round transcription assays were performed in transcription buffer (40 mM Tris pH 7.9, 10mM MgCl<sub>2</sub>, 0.6mM EDTA, 0.4mM KPO<sub>4</sub>, 1.5 mM DTT, 0.25 mg/mL BSA, 20% (v/v) glycerol). Reactions (15 µL volume) were performed by incubating *S. venezuelae* RNA polymerase (25 nM) with  $\sigma^{\text{HrdB}}$  (100 nM) for 5 minutes at 30 °C, followed by HelR<sub>SV</sub> and rifampin/RPP (in relevant reactions) with an additional 5-minute incubation. Template DNA (P<sub>AP3</sub>) was added to a final concentration of 10 nM and 2 µL of NTPs (0.4 mM ATP, GTP, CTP, and 0.1 mM UTP with 20 µCi[ $\alpha$ -<sup>32</sup>P] UTP (800 Ci/mmol)) were added to initiate the reaction which was allowed to proceed for 7.5 minutes at 30 °C. Reactions were stopped by the addition of an equal volume of loading buffer (7 M urea, 0.01% bromophenol blue, 20% glycerol) and analyzed on a 20% polyacrylamide gel containing 7 M urea. The bottom of the gel, containing unincorporated [ $\alpha$ -<sup>32</sup>P] UTP was manually excised with a scalpel and discarded. Transcripts were visualized using autoradiography with a Typhoon<sup>TM</sup> imager. Quantification of full length (run-off) transcripts was performed using densitometry with ImageJ (Schneider et al., 2012). Half-maximal inhibitory concentrations were determined with non-linear regression using Prism v7.05.

When demonstrating rifampin displacement *in-vitro* Arr-2 and its substrate NAD were used at concentrations of 5µM and 500µM, respectively. Reactions were performed as described above, except Arr and NAD were added following the 5-minute incubation with HelR<sub>SV</sub> and/or rifampin. These reactions were then initiated after different time

intervals with NTPs and template. Arr-2 was purified as described previously (Baysarowich et al., 2008).

### **Preparation of Template DNA**

P<sub>AP3</sub> DNA was synthesized a gBlock (IDT, **Figure S4**) with flanking *Hind* III and *Bam* HI sites and cloned into pUC19. The two nucleotides following the transcription start site were substituted with thymines, this enabled the detection of 2-3nt abortive products produced by RNAP inhibited by rifampin using [ $\alpha$ -<sup>32</sup>P] UTP. Sequence verified plasmid was used as the template for PCR using P<sub>AP3</sub> FP/RP to generate P<sub>AP3</sub>. PCR reactions were analyzed for purity on an agarose gel and purified using a GeneJET PCR purification kit (Thermo Fisher).

### **Rifamycin B purification**

*Amycolatopsis mediterranei* S699 was grown on solid Bennett's media for 7 days at 30°C, and 600 mL of YMG media (4 g glucose, 4 g yeast extract, 10 g malt extract in 1 L distilled water) was inoculated with ~10 single colonies worth of biomass (August et al., 1998). Following 7 days of growth (at 30 °C and 250 rpm), the production of rifamycin B was confirmed by LC-ESI-MS. The culture was mixed with an equal volume of methanol and returned to the shaker for 5 minutes. Cells and other particulates were removed by centrifugation (15 minutes, 6 000 x g), and the supernatant was decanted. Methanol was removed from the sample by rotary evaporation, and the remaining solution was lyophilized. Dried extract was resuspended in a 1:1 acetonitrile: H<sub>2</sub>O mixture (~5mL), centrifuged to remove insoluble material (8000 x g for 5 minutes). This crude extract was fractionated using reverse-phase flash chromatography (CombiFlash ISCO, using a

RediSep®RF Silver C18 26g column (Teledyne)) with a 20 minute linear gradient of water and acetonitrile (5 to 95% acetonitrile). Fractions containing high absorbance at 300 nm were checked for purity by LC-ESI-MS and fractions containing pure rifamycin B were pooled and lyophilized. Rifamycin B (~100mg) was obtained as fluffy yellow-orange powder, HRMS: Calculated for C<sub>39</sub>H<sub>49</sub>NO<sub>14</sub> 755.3153 Da, [M-H]<sup>-</sup> C<sub>39</sub>H<sub>49</sub>NO<sub>14</sub>: 754.3080 Da, found 754.3083 Da.<sup>1</sup>H and <sup>13</sup>C NMR performed in d<sub>6</sub>-DMSO are reported in **Table S3**.

## RPP SYNTHESIS

### Synthesis of N-[4(4-aminobenzoyl)phenyl]hex-5-ynamide (BPh-alkyne)

This compound was synthesized according to previously published methods (Salisbury and Cravatt, 2008), see **Figure S7**. 1 mL (8.9 mmol) of 5-hexynoic acid was added to a 50 mL round bottom flask followed by 15 mL DMF (anhydrous, Sigma Aldrich), 1.7 g (8.9 mmol) 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDAC, Sigma Aldrich), and 1.2 g (8.9 mmol) hydroxybenzotriazole (HOBt, Sigma-Aldrich). This solution was incubated for 5 minutes at room temperature with stirring. After the solution was clarified, 1.9 g (8.9 mmol) 4,4'-diaminobenzophenone (Sigma-Aldrich) was added, and the reaction (now a brown color) was incubated at room temperature for 24 hours with stirring. The reaction was applied to a reverse-phase CombiFlash ISCO (RediSep®RF Silver C18 26 g (Teledyne)) and purified using a water-acetonitrile linear gradient. Fractions were monitored by LC-ESI-MS, and pure fractions were pooled and lyophilized, 1.6g Bph-alkyne was obtained (60% yield).

### **Coupling of BPh-alkyne to rifamycin B to form RP-Alkyne**

85 mg (0.1125 mmol) of rifamycin B was dissolved in 1 mL of DMF in a 50 mL Falcon tube. Once rifamycin B had dissolved, 21.6 mg (0.1125 mmol) EDAC and 15.2 mg (0.1125 mmol) HOBt were added. This solution was incubated at room temperature with stirring for 20 minutes, and then 34.6 mg (0.1125 mmol) BPh-alkyne were added, and the reaction was allowed to proceed for 16 hours at room temperature. Monitoring of the reaction by LC-ESI-MS showed considerable starting product remaining after 16 hours, so 64.8 mg (0.3375 mmol) EDAC were added. The reaction was incubated for another 32 hours (48 hours total), after which no starting material could be observed by LC-ESI-MS. Reaction contents were applied directly to a 100mL Sephadex LH-20 column using a mobile phase of 100% methanol at a flow rate of ~2.5 mL/min, fractions were collected using a BioRad model 2110 fraction collector and analyzed by LC-ESI-MS. Fractions containing Rifamycin B-N-[4(4-aminobenzoyl)phenyl]hex-5-ynamide (RP-Alkyne) were further purified by reverse-phase flash chromatography using methodology previously described for rifamycin B, pure fractions were determined by LC-ESI-MS, pooled, and lyophilized, for a total of 41mg of RP-Alkyne (35% yield).

### **Huisgen cycloaddition to produce RPP**

4.5 mg (0.004 mmol) of RP-Alkyne was dissolved in 50  $\mu$ L of *t*-butanol, then added to 30  $\mu$ L H<sub>2</sub>O and vortexed briefly. 2.2mg (0.005 mmol) of Azide-PEG3-biotin (Sigma-Aldrich) was suspended in 10  $\mu$ L of DMSO and added to the *t*-butanol:H<sub>2</sub>O mixture, followed by 3  $\mu$ L sodium ascorbate (250 mg/mL in H<sub>2</sub>O) and 3  $\mu$ L of CuSO<sub>4</sub> (50 mg/mL in H<sub>2</sub>O), vortexing after each addition. Huisgen cycloaddition reaction was allowed to

proceed for 2 hours, after which RPP was purified by reverse phase flash chromatography as described previously using a RediSep®RF C18 4.3g column (Teledyne). Pure fractions were identified by LC-ESI-MS, pooled, and lyophilized to yield pure RPP (2 mg, 32% yield) as a fluffy yellow powder. HRMS: Calculated for  $C_{76}H_{97}N_9O_{20}S$  1487.6571 Da, [M-H]<sup>-</sup>  $C_{76}H_{96}N_9O_{20}S$ : 1486.6498 Da, found 1486.6499 Da. <sup>1</sup>H and <sup>13</sup>C NMR performed in d<sub>6</sub>-DMSO are reported in **Table S3**, see **Figure S7**.

### **RPP Photolabeling**

Photolabeling reactions were carried out at a 15µL scale in TBS in triplicate. *S. venezuelae* RNA polymerase, HelR<sub>SV</sub>, ATP were prepared as 15X stocks in TBS, and Rif and RPP were dissolved in DMSO as 15X stocks. 130 nM of RNA polymerase was incubated with HelR<sub>SV</sub>, ATP, and Rif where indicated for 5 minutes, after which RPP was added and allowed to incubate for another 5 minutes in total darkness. Parafilm was spread over a 96-well plate, and a gloved hand was used to create indentations in each well. Labeling reactions were aliquoted into these indentations, and the 96-well plate was placed under a UVP Blak-Ray B-100A lamp (365nm, Analytik Jena) approximately 20 cm from the bulb for 10 minutes. Reactions were mixed with 5µL of 4X Laemmli sample buffer (Bio-Rad), boiled for 5 minutes, separated on a 6% polyacrylamide gel, and transferred to PVDF membranes. Following transfer, blots were washed 3x with H<sub>2</sub>O for 2 minutes each and stained with Ponceau S (0.5% w/v in 1% acetic acid) for 1 minute. Blots were de-stained by 3x 30s washes with H<sub>2</sub>O and allowed to air dry before images for protein normalization were captured on a Chemidoc (BioRad). An additional 3x 5 minute washes with TBST were performed to remove all Ponceau S and the membranes were blocked,

washed, and imaged as described for  $\alpha$ -FLAG and  $\alpha$ -RpoB blots but were instead probed using a 1:25000 dilution of Streptavidin-HRP conjugate (Sigma Aldrich). Densitometry was used to quantify total protein (from Ponceau S) and signal from streptavidin-HRP using ImageJ (Schneider et al., 2012).

### **HelR/HelD network analysis**

This pipeline is publicly available (See **key resources table**) and is written in Snakemake v6.0.5 (Köster and Rahmann, 2012, 2018). To ensure a minimum consistent annotation for targeted genomes, all RefSeq assemblies from organisms classified according to the NCBI taxonomy in the phyla Actinobacteria and Firmicutes were downloaded using the NCBI-genome-download v0.3.0 (<https://github.com/kblin/ncbi-genome-download>), totaling 7457 genomes as of Jan. 12<sup>th</sup>, 2021. Up to 500bp upstream of each coding sequence feature from each of the downloaded genomes was extracted into separate fasta files. The previously published RAE sequences were used to generate an alignment and nucleotide Hidden Markov Model (HMM) using nhmmer from the HMMER 3.1b software (Eddy, 2011; Spanogiannopoulos et al., 2014). This RAE HMM was used to search every upstream sequence of every CDS of every downloaded genome. The search was limited to the top strand only, and putative RAE sequences were identified as having a minimum length of 15 bp and a minimum score of 10.0 bits.

The HelR sequence from *M. smegmatis* (WP\_003893549.1) was used to search the NCBI nr database using BLASTp with default parameters (Oyama et al., 2008). The top 5000 sequence hits (by E-value) were downloaded as a fasta file, sorted by length, and subjected to a clustering step using usearch v11.0.667 using a threshold of 99% identity

(Edgar, 2010). This resulted in a set of 4906 clusters. The centroid sequence from each cluster was used to query the downloaded genome assemblies using an E-value cutoff of  $1e-255$  using DIAMOND v 0.9.14.115 (Buchfink et al., 2014). This resulted in a set of 15,136 putative HelD sequences. These sequences were annotated with the RAE sequence, if identified, in their upstream sequences.

The RAE annotated HelD sequences from RefSeq Actinobacteria, and Firmicutes were then sorted by length and subjected to a cluster analysis using usearch at a threshold of 50% identity, resulting in 417 clusters. A cursory glance at these clusters revealed that several were entirely composed of rifamycin glycosyltransferases (Rgt), a RAE associated resistance enzyme. These were included in our dataset because of predicted Rgt-HelR fusions. We have previously noted the presence of a single RAE found upstream of *rgt* and *helR* in what is presumably a single operon, in various *Streptomyces* genomes. Because the number of Rgt-HelR fusions scant ( $n = 2$ ) we believed these to be errors arising from sequencing/assembly of these genomes and are not likely to be true fusion proteins. Because these fusion proteins were unintentionally used to query Refseq genomes, they pulled in many hits for Rgt, explaining the presence of glycosyltransferase clusters in our data. We felt it appropriate to remove these clusters from our main analysis, although they are still reported in **Supplementary Item 2**. To do this we used CDsearch to identify conserved domains from the centroid sequence of each cluster (**Supplementary Item 2**) as expected the vast majority of clusters possess a complete or partial HelD domain, and the glycosyltransferases could be easily removed (Lu et al., 2020). After excluding Rgts a

total of 29 protein clusters were identified as having at least one member sequence possessing an upstream RAE (Edgar, 2010).

The 3988 RAE-associated HelD (putative HelR) sequences were aligned using mafft v7.310 (using the “—auto” parameter), resulting in a highly gapped alignment of 3874 columns (Kato and Standley, 2013). This alignment was trimmed using TrimAl v1.4.rev22 (using the “--automated1” parameter) to an alignment of 357 columns (Capella-Gutiérrez et al., 2009).

To generate the SSN, the combined set of sequences identified in the RefSeq genomes using the HelD search set were subjected to an all-vs-all search using DIAMOND (Buchfink et al., 2014) with an E-value cutoff of  $1 \times 10^{-10}$ . Self-hits were filtered out of these results. The network was further refined by using a minimum edge score cutoff of 700 bits for each similarity result to produce the final network. The network was visualized in cytoscape (Shannon et al., 2003).

### **Quantification and Statistical analysis**

Statistical details such as the number of replicates and tests for significance used which are not stated in the main text can be found in the figure legends and Method Details sections.

## REFERENCES

- Adams, K.N., Takaki, K., Connolly, L.E., Wiedenhof, H., Winglee, K., Humbert, O., Edelstein, P.H., Cosma, C.L., and Ramakrishnan, L. (2011). Drug tolerance in replicating mycobacteria mediated by a macrophage-induced efflux mechanism. *Cell* *145*, 39–53. <https://doi.org/10.1016/j.cell.2011.02.022>.
- August, P.R., Tang, L., Yoon, Y.J., Ning, S., Rolf, M., Yu, T., Taylor, M., Hoffmann, D., Kim, C., Zhang, X., et al. (1998). Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* *5*, 69–79. [https://doi.org/10.1016/s1074-5521\(98\)90141-7](https://doi.org/10.1016/s1074-5521(98)90141-7).
- Van Bambeke, F., Van Laethem, Y., Courvalin, P., and Tulkens, P.M. (2004). Glycopeptide antibiotics: From conventional molecules to new derivatives. *Drugs* *64*, 913–936. <https://doi.org/10.2165/00003495-200464090-00001>.
- Baysarowich, J., Koteva, K., Hughes, D.W., Ejim, L., Griffiths, E., Zhang, K., Junop, M., and Wright, G.D. (2008). Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4886–4891. <https://doi.org/10.1073/pnas.0711939105>.
- Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat. Rev. Microbiol.* *14*, 320–330. <https://doi.org/10.1038/nrmicro.2016.34>.
- Brown, E.D., and Wright, G.D. (2016). Antibacterial drug discovery in the resistance era. *Nature* *529*, 336–343. <https://doi.org/10.1038/nature17042>.
- Buchfink, B., Xie, C., and Huson, D.H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* *12*, 59–60. <https://doi.org/10.1038/nmeth.3176>.
- Burgess, R.R., and Jendrisak, J.J. (1975). Procedure for the Rapid, Large-Scale Purification of *Escherichia coli* DNA-Dependent RNA Polymerase Involving Polymin P Precipitation and DNA-Cellulose Chromatography. *Biochemistry* *14*, 4634–4638. <https://doi.org/10.1021/bi00692a011>.
- Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001). Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* *104*, 901–912. [https://doi.org/10.1016/s0092-8674\(01\)00286-0](https://doi.org/10.1016/s0092-8674(01)00286-0).
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* *25*, 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Chen, J., Boyaci, H., and Campbell, E.A. (2021). Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* *19*, 95–109. <https://doi.org/10.1038/s41579-020-00450-2>.
- Copp, J.N., Akiva, E., Babbitt, P.C., and Tokuriki, N. (2018). Revealing Unexplored Sequence-Function Space Using Sequence Similarity Networks. *Biochemistry* *57*, 4651–4662. <https://doi.org/10.1021/acs.biochem.8b00473>.
- Crowe-McAuliffe, C., Graf, M., Huter, P., Takada, H., Abdelshahid, M., Nováček, J., Murina, V., Atkinson, G.C., Haurlyuk, V., and Wilson, D.N. (2018). Structural basis for antibiotic resistance mediated by the *Bacillus subtilis* ABCF ATPase VmlR. *Proc. Natl. Acad. Sci. U. S. A.* *115*, 8978–8983. <https://doi.org/10.1073/pnas.1808535115>.
- Dabbs, E. (1987). Rifampicin inactivation by *Rhodococcus* and *Mycobacterium* species. *FEMS Microbiol. Lett.* *44*, 395–399. <https://doi.org/10.1111/j.1574-6968.1987.tb02320.x>.

- Dabbs, E.R., Yazawa, K., Mikami, Y., Miyaji, M., Morisaki, N., Iwasaki, S., and Furihata, K. (1995). Ribosylation by mycobacterial strains as a new mechanism of rifampin inactivation. *Antimicrob. Agents Chemother.* *39*, 1007–1009. <https://doi.org/10.1128/AAC.39.4.1007>.
- Dangkulwanich, M., Ishibashi, T., Bintu, L., and Bustamante, C. (2014). Molecular Mechanisms of Transcription through Single-Molecule Experiments. *Chem. Rev.* *114*, 3203–3223. <https://doi.org/10.1021/cr400730x>.
- Dillingham, M.S., and Dillingham, M. (2011). Colworth Medal Lecture: Superfamily I helicases as modular components of DNA-processing machines. *Biochem. Soc. Trans.* *39*, 413–423. <https://doi.org/10.1042/BST0390413>.
- Dönhöfer, A., Franckenberg, S., Wickles, S., Berninghausen, O., Beckmann, R., and Wilson, D.N. (2012). Structural basis for TetM-mediated tetracycline resistance. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 16900–16905. <https://doi.org/10.1073/pnas.1208037109>.
- Dürckheimer, W. (1975). Tetracyclines: Chemistry, Biochemistry, and Structure-Activity Relations. *Angew. Chemie Int. Ed. English* *14*, 721–734. <https://doi.org/10.1002/anie.197507211>.
- Eddy, S.R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* *7*. <https://doi.org/10.1371/journal.pcbi.1002195>.
- Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* *26*, 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
- Epshtein, V., Kamarthapu, V., Mcgary, K., Svetlov, V., Proshkin, S., Mironov, A., and Nudler, E. (2015). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* *505*, 372–377. <https://doi.org/10.1038/nature12928.UvrD>.
- Fairman-Williams, M.E., Guenther, U.P., and Jankowsky, E. (2010). SF1 and SF2 helicases: Family matters. *Curr. Opin. Struct. Biol.* *20*, 313–324. <https://doi.org/10.1016/j.sbi.2010.03.011>.
- Floss, H.G., and Yu, T.-W. (2005). Rifamycin Mode of Action, Resistance, and Biosynthesis. *Chem. Rev.* *105*, 621–632. <https://doi.org/10.1021/CR030112J>.
- Gillespie, S. (2002). Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrob. Agents Chemother.* *46*, 267–274. <https://doi.org/10.1128/AAC.46.2.267>.
- Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci.* *100*, 1541–1546. <https://doi.org/10.1073/pnas.0337542100>.
- Hong, H.J., Hutchings, M.I., Hill, L.M., and Buttner, M.J. (2005). The role of the novel fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J. Biol. Chem.* *280*, 13055–13061. <https://doi.org/10.1074/jbc.M413801200>.
- Hurst-Hess, K., Biswas, R., Yang, Y., Rudra, P., Lasek-Nesselquist, E., and Ghosh, P. (2019). Mycobacterial SigA and SigB cotranscribe essential housekeeping genes during exponential growth. *MBio* *10*, 1–17. <https://doi.org/10.1128/mBio.00273-19>.
- Hutter, B., Fischer, C., Jacobi, A., Schaab, C., and Loferer, H. (2004). Panel of *Bacillus subtilis* Reporter Strains Indicative of Various Modes of Action. *Antimicrob. Agents Chemother.* *48*, 2588–2594. <https://doi.org/10.1128/AAC.48.7.2588>.
- Javid, B., Sorrentino, F., Toosky, M., Zheng, W., Pinkham, J.T., Jain, N., Pan, M., Deighan, P., and Rubin, E.J. (2014). Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 1132–1137. <https://doi.org/10.1073/pnas.1317580111>.

- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* *30*, 772–780. <https://doi.org/10.1093/molbev/mst010>.
- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K., and Hopwood, D. (2000). *Practical Streptomyces Genetics* (Colney, Norwich NR4 7UH, England: John Innes Foundation).
- Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie - Int. Ed.* *40*, 2004–2021. [https://doi.org/10.1002/1521-3773\(20010601\)40:11<2004::AID-ANIE2004>3.0.CO;2-5](https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5).
- Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M., et al. (2017). Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst.* *4*, 291-305.e7. <https://doi.org/10.1016/j.cels.2016.12.013>.
- Köster, J., and Rahmann, S. (2012). Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* *28*, 2520–2522. <https://doi.org/10.1093/bioinformatics/bts480>.
- Köster, J., and Rahmann, S. (2018). Corrigendum: Snakemake - A scalable bioinformatics workflow engine (*Bioinformatics* (2012) 28:19 (2520-2522) DOI: 10.1093/bioinformatics/bts480). *Bioinformatics* *34*, 3600. <https://doi.org/10.1093/bioinformatics/bty350>.
- Koteva, K., Cox, G., Kelso, J.K., Surette, M.D., Zubyk, H.L., Ejim, L., Stogios, P., Savchenko, A., Sørensen, D., and Wright, G.D. (2018). Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem. Biol.* *25*, 403–412.e5. <https://doi.org/10.1016/j.chembiol.2018.01.009>.
- Kouba, T., Koval', T., Sudzinová, P., Pospíšil, J., Brezovská, B., Hnilicová, J., Šanderová, H., Janoušková, M., Šiková, M., Halada, P., et al. (2020). Mycobacterial HelD is a nucleic acids-clearing factor for RNA polymerase. *Nat. Commun.* *11*, 1–13. <https://doi.org/10.1038/s41467-020-20158-4>.
- Li, W., Atkinson, G.C., Thakor, N.S., Allas, U., Lu, C.C., Yan Chan, K., Tenson, T., Schulten, K., Wilson, K.S., Haurlyiuk, V., et al. (2013). Mechanism of tetracycline resistance by ribosomal protection protein Tet(O). *Nat. Commun.* *4*. <https://doi.org/10.1038/ncomms2470>.
- Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M., et al. (2018). Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Mol. Cell* *70*, 60–71.e15. <https://doi.org/10.1016/j.molcel.2018.02.026>.
- Lobritz, M.A., Belenky, P., Porter, C.B.M., Gutierrez, A., Yang, J.H., Schwarz, E.G., Dwyer, D.J., Khalil, A.S., and Collins, J.J. (2015). Antibiotic efficacy is linked to bacterial cellular respiration. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 8173–8180. <https://doi.org/10.1073/pnas.1509743112>.
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., et al. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* *48*, D265–D268. <https://doi.org/10.1093/nar/gkz991>.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., et al. (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* *47*, D351–D360. <https://doi.org/10.1093/nar/gky1100>.
- Mosaei, H., Molodtsov, V., Kepplinger, B., Harbottle, J., Moon, C.W., Jeeves, R.E., Ceccaroni, L., Shin, Y., Morton-Laing, S., Marrs, E.C.L., et al. (2018). Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant *Mycobacterium tuberculosis*. *Mol. Cell* *72*, 263–274.e5. <https://doi.org/10.1016/j.molcel.2018.08.028>.
- Murale, D.P., Hong, S.C., Haque, M.M., and Lee, J.S. (2017). Photo-affinity labeling (PAL) in chemical proteomics: A handy tool to investigate protein-protein interactions (PPIs). *Proteome Sci.* *15*, 1–34.

<https://doi.org/10.1186/s12953-017-0123-3>.

Myronovskyi, M., Welle, E., Fedorenko, V., and Luzhetskyy, A. (2011).  $\beta$ -Glucuronidase as a Sensitive and Versatile Reporter in Actinomycetes. *Appl. Environ. Microbiol.* *77*, 5370–5383.

<https://doi.org/10.1128/AEM.00434-11>.

Newing, T.P., Oakley, A.J., Miller, M., Dawson, C.J., Brown, S.H.J., Bouwer, J.C., Tolun, G., and Lewis, P.J. (2020). Molecular basis for RNA polymerase-dependent transcription complex recycling by the helicase-like motor protein HelD. *Nat. Commun.* *11*, 1–11. <https://doi.org/10.1038/s41467-020-20157-5>.

Oyama, R.K., Clauss, M.J., Formanová, N., Kroymann, J., Schmid, K.J., Vogel, H., Weniger, K., Windsor, A.J., and Mitchell-Olds, T. (2008). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* *25*, 3389–3402.

<https://doi.org/10.1093/oxfordjournals.molbev.a026201>.

Pei, H.H., Hilal, T., Chen, Z.A., Huang, Y.H., Gao, Y., Said, N., Loll, B., Rappsilber, J., Belogurov, G.A., Artsimovitch, I., et al. (2020). The  $\delta$  subunit and NTPase HelD institute a two-pronged mechanism for RNA polymerase recycling. *Nat. Commun.* *11*, 1–14. <https://doi.org/10.1038/s41467-020-20159-3>.

Piccaro, G., Pietraforte, D., Giannoni, F., Mustazzolu, A., and Fattorini, L. (2014). Rifampin induces hydroxyl radical formation in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* *58*, 7527–7533. <https://doi.org/10.1128/AAC.03169-14>.

Quan, S., Venter, H., and Dabbs, E.R. (1997). Ribosylative Inactivation of Rifampin by *Mycobacterium smegmatis* Is a Principal Contributor to Its Low Susceptibility to This Antibiotic. *Antimicrob. Agents Chemother.* *41*, 2456–2460. .

Ramaswamy, S., and Musser, J.M. (1998). Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber. Lung Dis.* *79*, 3–29.

<https://doi.org/10.1054/tuld.1998.0002>.

Raney, K.D., Byrd, A.K., and Aarattuthodiyil, S. (2013). Structure and mechanisms of SF1 DNA helicases. *Adv. Exp. Med. Biol.* *767*, 17–46. [https://doi.org/10.1007/978-1-4614-5037-5\\_2](https://doi.org/10.1007/978-1-4614-5037-5_2).

Salisbury, C.M., and Cravatt, B.F. (2008). Optimization of Activity-Based Probes for Proteomic Profiling of Histone Deacetylase Complexes. *JACS* *130*, 2184–2194. <https://doi.org/10.1021/ja074138u>.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* *9*, 671–675. <https://doi.org/10.1038/nmeth.2089>.

Sensi, P. (1983). History of the Development of Rifampin. *Rev. Infect. Dis.* *5*, S402–S406. .

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Nada, A., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models. *Genome Res.* *13*, 426. <https://doi.org/10.1101/gr.1239303.metabolite>.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annu. Rev. Biochem.* *72*, 23–50. <https://doi.org/10.1146/annurev.biochem.76.052305.115300>.

Spanogiannopoulos, P. (2014). Exploring Rifamycin Inactivation from the Soil Microbiome. Doctoral Thesis. McMaster University.

Spanogiannopoulos, P., Thaker, M., Koteva, K., Waglechner, N., and Wright, G.D. (2012). Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob. Agents Chemother.* *56*, 5061–5069. <https://doi.org/10.1128/AAC.01166-12>.

Spanogiannopoulos, P., Waglechner, N., Koteva, K., and Wright, G.D. (2014). A rifamycin inactivating

phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 7102–7107. <https://doi.org/10.1073/pnas.1402358111>.

Su, W., Kumar, V., Ding, Y., Ero, R., Serra, A., Lee, B.S.T., Wong, A.S.W., Shi, J., Sze, S.K., Yang, L., et al. (2018). Ribosome protection by antibiotic resistance ATP-binding cassette protein. *Proc. Natl. Acad. Sci. U. S. A.* *115*, 5157–5162. <https://doi.org/10.1073/pnas.1803313115>.

Surette, M., and Wright, G.D. (2017). Lessons from the Environmental Antibiotic Resistome. *Annu. Rev. Microbiol.* *71*, 309–329. <https://doi.org/10.1146/annurev-micro-090816-093420>.

Surette, M.D., Spanogiannopoulos, P., and Wright, G.D. (2021). The Enzymes of the Rifamycin Antibiotic Resistome. *Acc. Chem. Res.* *111*. <https://doi.org/10.1021/acs.accounts.1c00048>.

Svetlov, M.S., Vázquez-Laslop, N., and Mankin, A.S. (2017). Kinetics of drug-ribosome interactions defines the cidalty of macrolide antibiotics. *Proc. Natl. Acad. Sci. U. S. A.* *114*, 13673–13678. <https://doi.org/10.1073/pnas.1717168115>.

Tang, X., Li, J., Millán-Aguiñaga, N., Zhang, J.J., O’Neill, E.C., Ugalde, J.A., Jensen, P., Mantovani, S.M., and Moore, B.S. (2015). Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-direct Genome Mining. *ACS Chem. Biol.* *10*, 2841–2849. <https://doi.org/10.1021/acscchembio.5b00658>. Identification.

Wehrli, W., Knüsel, F., Schmid, K., and Staehelin, M. (1968). Interaction of rifamycin with bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* *61*, 667–673. <https://doi.org/10.1073/pnas.61.2.667>.

Weinstein, M.P., Patel, J.B., Burnhman, C.-A., and Zimmer, B.L. (2018). CLSI. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically* 11th Edition (Wayne, PA: Clinical and Laboratory Standards Institute).

World Health Organization (2008). *Anti-Tuberculosis Drug Resistance in the World*.

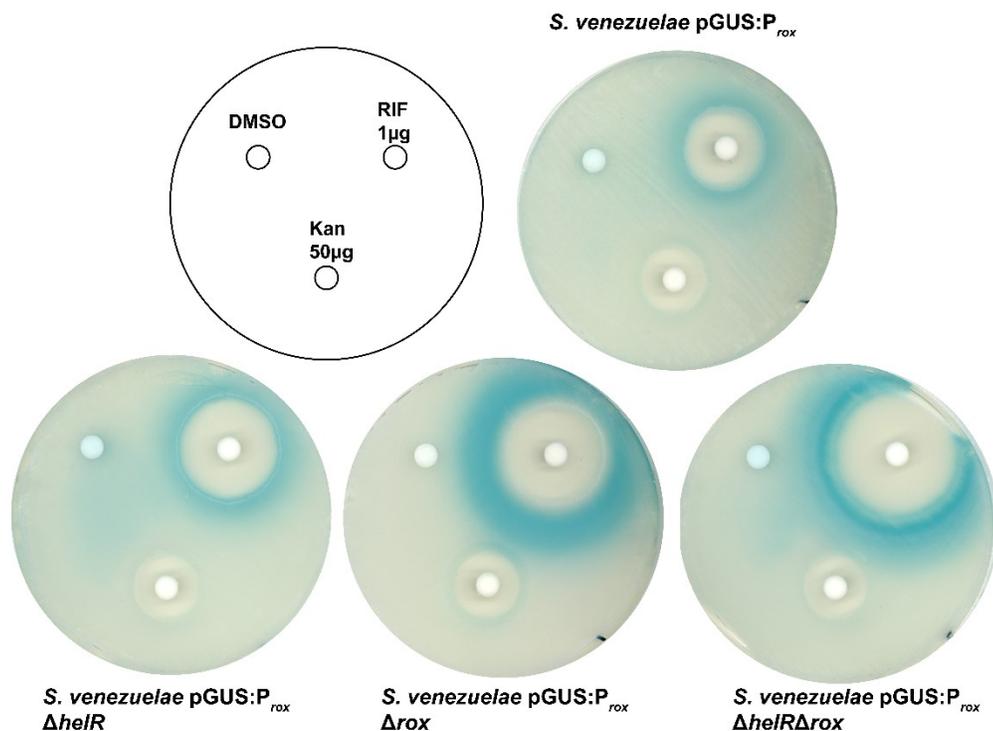
Wiedermannová, J., Sudzinová, P., Koval’, T., Rabatinová, A., Šanderova, H., Ramaniuk, O., Rittch, Š., Dohnálek, J., Fu, Z., Halada, P., et al. (2014). Characterization of HelD, an interacting partner of RNA polymerase from *Bacillus subtilis*. *Nucleic Acids Res.* *42*, 5151–5163. <https://doi.org/10.1093/nar/gku113>.

Wilson, D.N., Hauryliuk, V., Atkinson, G.C., and O’Neill, A.J. (2020). Target protection as a key antibiotic resistance mechanism. *Nat. Rev. Microbiol.* *18*, 637–648. <https://doi.org/10.1038/s41579-020-0386-z>.

Yam, W.C., Tam, C.M., Leung, C.C., Tong, H.L., Chan, K.H., Leung, E.T.Y., and Wong, K.C. (2004). Direct Detection of Rifampin-Resistant *Mycobacterium tuberculosis* in Respiratory Specimens by PCR-DNA Sequencing. *J. Clin. Microbiol.* *42*, 4438–4443. <https://doi.org/10.1128/JCM.42.10.4438>.

Zhu, J.H., Wang, B.W., Pan, M., Zeng, Y.N., Rego, H., and Javid, B. (2018). Rifampicin can induce antibiotic tolerance in mycobacteria via paradoxical changes in rpoB transcription. *Nat. Commun.* *9*. <https://doi.org/10.1038/s41467-018-06667-3>.

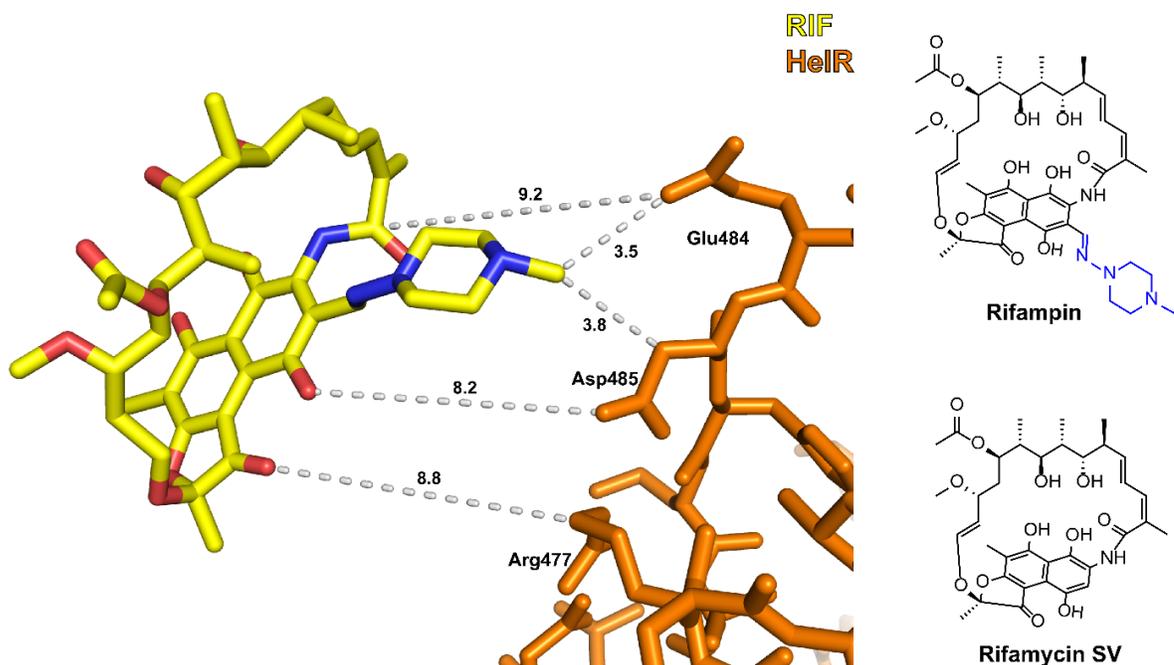
### Supplementary Material



**Figure S1. Induction through the RAE does not require HelR, Related to Figure 1** *S. venezuelae* mutants transformed with pGUS:P<sub>rox</sub> were streaked for confluence on media containing colorimetric GUS substrate. Zones of blue indicate GUS activity and therefore induction of P<sub>rox</sub>. Although some strains are more susceptible than others all are induced by sub-inhibitory rifampin. *S. venezuelae*  $\Delta rox$  and  $\Delta rox\Delta helR$  appear to have the higher expression in response to rifampin. We hypothesize this is because cells that possess *rox* can inactivate rifampin, dampening induction caused by these compounds over time.

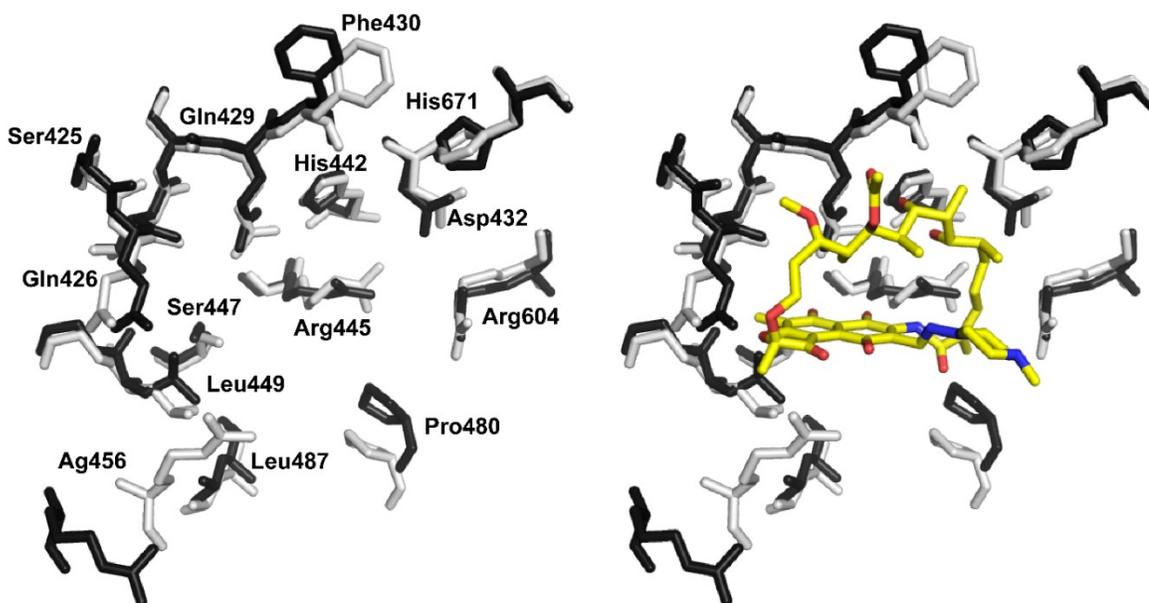
**Table S1** Top 10 Enriched and Depleted RNAP associated protein,  
**Related to Figure 3A**

<b>Protein</b>	<b>Annotation / Predicted function</b>	<b>Fold Enrichment</b>
HelR	Superfamily 1 helicase	426
SVEN_4182	TetR Family transcriptional regulator	102
RplQ	Ribosomal protein L17	79
SVEN_5967	Butyryl-CoA dehydrogenase	40
SVEN_2381	DUF2344 domain-containing protein	19
SVEN_3586	Pyruvate dehydrogenase E1 beta subunit	17
SVEN_6463	Outer membrane protein RomA	15
GitX	Glutamate tRNA ligase	11
SVEN_5919	Ferredoxin sulfite reductase	10
SVEN_1435	Putative ABC transporter ATP-binding protein	8
<b>Protein</b>	<b>Annotation / Predicted function</b>	<b>Fold Depletion</b>
SVEN_0613	Secreted protein (Arabinose binding)	19.2
SVEN_7112	SAM-dependent methyltransferase	14.4
SVEN_4686	Secreted protein (OmpA-like domain)	8.5
GrpE	Nucleotide exchange factor for DnaK	5.3
RpoZ	RNA polymerase $\omega$ subunit	5.3
SVEN_1812	Cytochrome C oxidase subunit 1	4.5
AlaS	Alanine tRNA ligase	4.1
SVEN_1149	Enoyl ACP reductase	3.6
SVEN_3043	Pyruvate formate-lyase	3.6
SVEN_2206	PucR family regulator	3.2



**Supplementary Figure 2. Distance between HelR<sub>M<sub>s</sub></sub> PCh loop residues and rifampin, related to Figure 3D. (Left)** Closest contact between the PCh loop and rifampin occurs on the methyl group of the N-methylpiperazine synthetic tail. This is unlikely to drive displacement as an actual clash is not observed and HelR<sub>S<sub>V</sub></sub> confers resistance to rifamycins which lack this tail entirely (rifamycin SV). Distance from PCh loop residues to core elements of rifamycins, what would be needed to clash with rifamycin SV, are considerably larger (8.2-9.2Å). **(right)** Structure of rifampin and rifamycin SV, semisynthetic tail of rifampin is colored blue. This figure was produced using an alignment of RpoB and Rifampin from 6CCV and RNAP:HelR<sub>M<sub>s</sub></sub> complex in state II (6YYS).

RNAP-RIF complex  
RNAP-HeIR complex State II

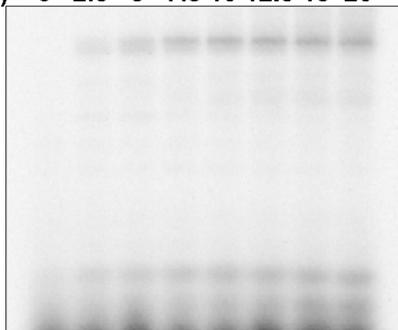


**Supplementary Figure 3. Rifamycin binding pocket in a State II RNAP:HeIR<sub>M5</sub> complex, related to Figure 3D.** Overlay of residues making up the rifamycin binding pocket from Rifampin bound RpoB (grey) and RNAP in complex with HeIR (during state II, black). We observed substantial conformational change only for Arg456 which moves 7.5Å due to an interaction with the PCh loop. However, the loss of the non-polar/Van der waals interactions from this residue alone are unlikely to eject rifamycins from their binding site. At the 3.08Å resolution of 6YYS, subtle changes in key residues (such as hydrogen partners His442, Ser447, and Arg445) cannot be unequivocally excluded.

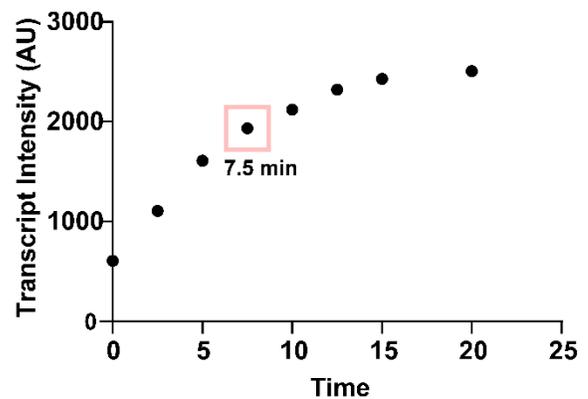
**P**<sub>AP3</sub>

5' -AAGCTTATCTATGGATGACCGAACCTGGTCTT**TGACT**CCATTGCCGGATTTGTATT**TAGACT**GGCAGGGTTGCC  
 CGAAGAACCATAAGATCCATTAAGGGCGAAGTACAAAGCTT-3'

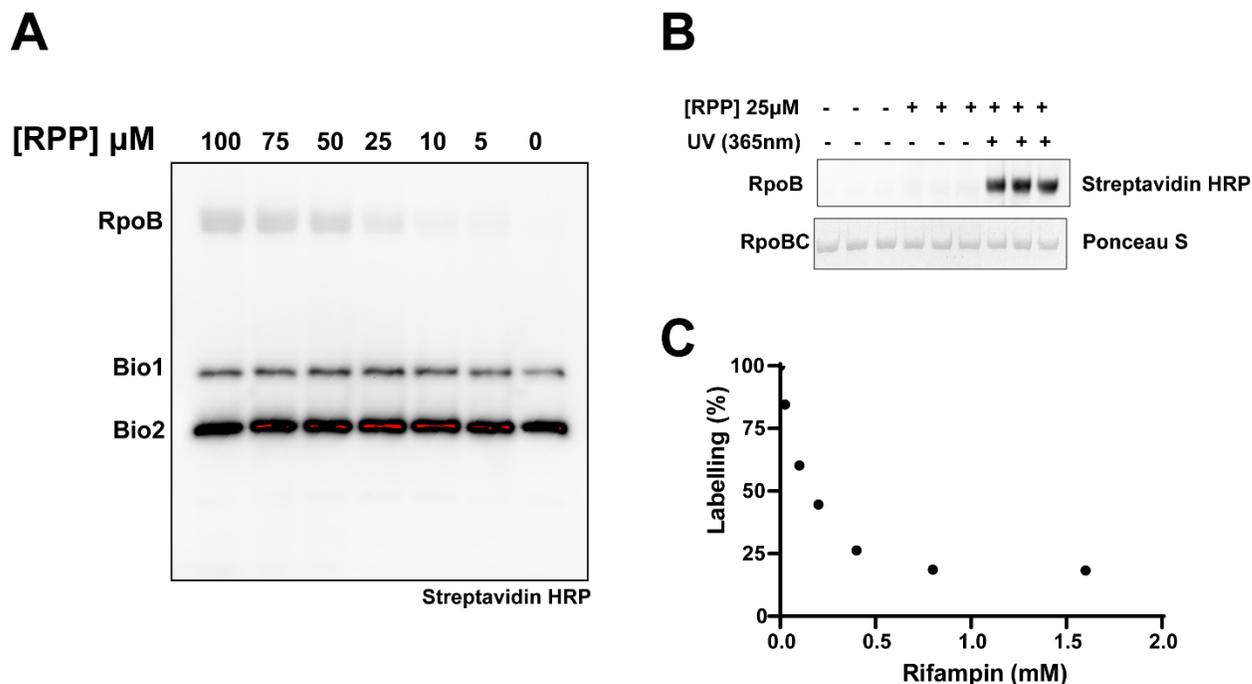
Time(min) 0 2.5 5 7.5 10 12.5 15 20



run off  
transcript



**Supplementary Figure 4. Progress curve for multiple round *in-vitro* transcription using P<sub>AP3</sub>, related to Figure 3F-H.** The AP3 promoter (top) is a strong constitutive promoter that controls expression of ribosomal RNA in *Mycobacterium tuberculosis*. (left) 20% Urea-PAGE analysis of transcription reactions of various lengths. (right) Quantified transcript; the 7.5-minute reaction was used for all *in vitro* transcription reactions in this manuscript as it is at the end of the linear portion of the reaction.

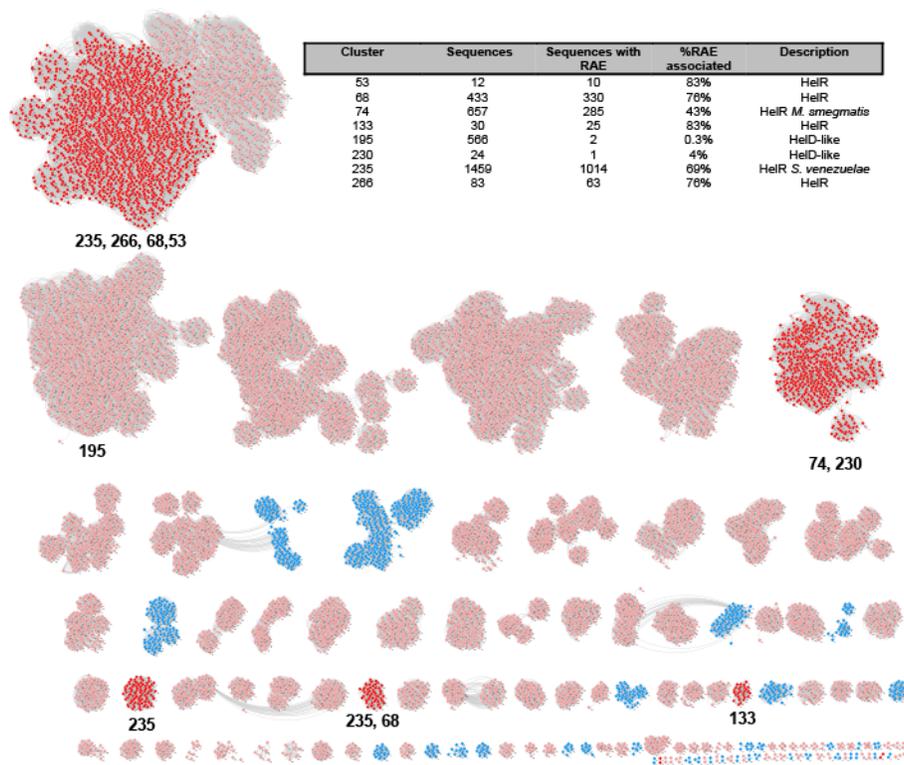


**Supplementary Figure 5. RPP labelling controls, related to Figure 4D-G.** **A) Endogenous biotinylated proteins** The full blot from Figure 4D. RNAP was incubated with RPP at various concentrations and crosslinked using long wave UV light and labelling was detected by western blotting with streptavidin-HRP. Bio1 and Bio2 denote endogenously biotinylated proteins in our RNAP preparations. Fortunately, these proteins migrated at  $\sim 55$  and  $\sim 70$  kDa and therefore did not interfere with our detection of labelled RpoB. **B) Labelling by RPP is UV dependent** Western blotting of RPP labelling reactions using streptavidin HRP. Shown below is the Ponceau S signal for RpoBC, used to ensure equal loading. Each lane represents an independent replicate. **C) Rifampin blocks labelling by RPP** Densitometric quantification of the experiment depicted in Figure 4E, demonstrating competition between rifampin and RPP. Labelling was expressed as a percentage relative to the 0mM Rifampin control.

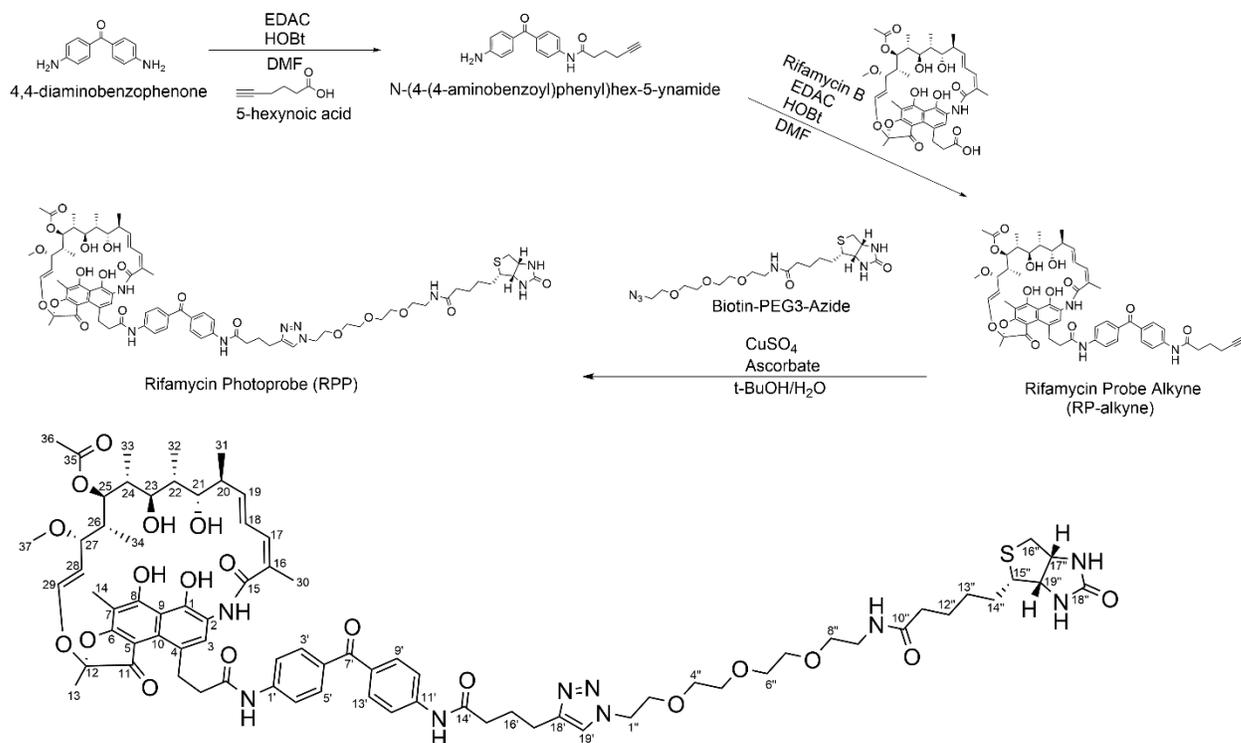
**Table S2** Clusters with  $\geq 1$  member associated with a RAE, related to **Figure 5**

Cluster No.	No. Sequences	No. Sequences with RAE	Fraction with RAE	Description
8	1	1	1	Gtf-HelD fusion
12	1	1	1	Gtf-HelD fusion
53	12	10	0.83	HelD
68	433	330	0.76	HelD
74	657	285	0.43	HelD
124	1	1	1	HelD
133	30	25	0.83	HelD
195	566	2	0	HelD
227	3	3	1	HelD_partial
230	24	1	0.04	HelD
235	1459	1014	0.69	HelD
239	5	2	0.4	HelD_partial
243	1	1	1	HelD_partial
247	1	1	1	HelD_partial
258	1	1	1	HelD_partial
259	1	1	1	HelD_partial
266	83	63	0.76	HelD
270	3	3	1	HelD_partial
280	1	1	1	HelD_partial
313	1	1	1	HelD_partial
319	3	2	0.67	HelD_partial
342	3	1	0.33	HelD_partial
346	82	36	0.44	Gtf
354	7	6	0.86	HelD_partial
361	237	145	0.61	Gtf
367	332	134	0.4	Gtf
370	23	7	0.3	Gtf
381	7	1	0.14	HelD_partial
384	5	5	1	HelD_partial
391	1	1	1	HelD_partial
405	2	2	1	HelD_partial
414	1	1	1	HelD_partial
417	1	1	1	HelD_partial

Gtf = Glycosyltransferase



**Supplementary Figure 6. Distribution of HelR clusters in the Sequence Similarity Network of HelD-like proteins, related to Figure 5.** The SSN from Figure 11 with the protein cluster identities associated with RAEs noted on the network.



**Supplementary Figure 7 Rifamycin Photoprobe (RPP) synthetic route and NMR assignment key, related to STAR methods.** (Top) Scheme for RPP synthesis, see STAR methods for more details. (Bottom) Structure of RPP with carbons labelled, accompanies the <sup>1</sup>H and <sup>13</sup>C NMR assignments reported in **Table S3**.

**Table S3** Chemical shifts for Rifamycin B and Rifampin Photo Probe in dms<sub>o</sub>-d<sub>6</sub>, reported in ppm. **Related to STAR Methods.**

#	<sup>1</sup> H Rif B	<sup>13</sup> C Rif B	<sup>1</sup> H RPP	<sup>13</sup> C RPP
1	-	145.17	-	146.17
2	-	119.37	-	119.05
3	7.39 (s, 1H)	108.71	7.67 (s, 1H)	107.62
4	-	145.34	-	142.60
5	-	114.07	-	114.51
6	-	171.96	-	169.60
7	-	100.26	-	99.51
8	-	114.07	-	114.51
9	-	115.88	-	114.51
10	-	98.72	-	99.83 (HMBC)
11	-	184.53	-	184.37 (HMBC)
12	-	108.73	-	107.62 (HMBC)
13	1.60 (s, 3H)	22.14	1.68 (s, 3H)	22.36
14	2.01 (s, 3H)	7.16	1.95 (s, 3H)	7.48
15	-	168.88	-	169.70
16	-	130.85	-	130.88
17	6.19 (m, 1H)	131.34	6.14 (m, 1H)	130.81
18	6.35 (d, J = 14.5 Hz, 1H)	124.6	6.44 (m, 1H)	122.30
19	5.91 (s, 1H)	139.17	5.95 (m, 1H)	138.45
20	2.14 (m, 1H)	37.58	2.15 (m, 1H)	38.42
21	3.68 (s, 1H)	71.64	3.69 (m, 1H)	73.23
22	1.56(m, 1H)	32.25	1.59 (m, 1H)	33.4
23	2.88 – 2.82 (m, 1H)	75.88	2.80 (m, 1H)	75.95
24	1.22 (m, 1H)	37.48	1.26 (m, 1H)	37.65
25	4.99 (d, J = 10.7 Hz, 1H)	72.9	5.01(m, 1H)	73.35
26	0.88 (m, 1H)	39.47	0.88 (m, 1H)	40.21
27	3.21 (m, 1H)	76.04	3.17 (m, 1H)	75.41
28	4.89 (s, 1H)	117.65	4.93 (m, 1H)	117.2
29	6.16 (m, 1H)	142.52	6.18 (m, 1H)	143.10
30	1.94 (s, 3H)	20.70	1.97 (m, 3H)	20.59
31	0.83 (m, 3H)	17.58	0.87 (m, 3H)	18.17
32	0.82 (m, 3H)	11.17	0.86 (m, 3H)	12.14
33	0.40 (s, 3H)	8.05	0.32 (m, 3H)	8.22
34	-0.45	9.03	-0.26 (m, 3H)	8.69
35	-	172.87	-	172.08
36	1.92 (s, 3H)	19.63	1.93 (m, 3H)	20.01
37	2.84 (s, 3H)	55.71	2.88 (m, 3H)	55.40
-CH <sub>2</sub> -COOH	4.55 (d, J = 13.1 Hz, 2H)	66.53	4.47 (t, J = 5.3 Hz, 2H)	49.20
-CH <sub>2</sub> -CO- NH	9.57 (s, 1H)	-	- <b>10.2*</b> (s, 1H)	171.53 -

1'	-	-	-	146.18
2'/6'	-	-	8.15 (m, 1H)	118.99
3'/5'	-	-	7.73 (m, 1H)	130.91
4'	-	-	-	131.64
7'	-	-	-	193.46
8'	-	-	-	132.29
9'/13'	-	-	7.79 (m, 1H)	130.67
10'/12'	-	-	7.78 (m, 1H)	118.17
11'	-	-	-	146.18
14'	-	-	-	171.53
15'	-	-	2.68 (t, J = 7.6 Hz, 2H)	24.56
16'	-	-	1.96 (m, 2H)	24.82
17'	-	-	2.44 (t, J = 7.4 Hz, 2H)	35.89
18'	-	-	-	nd
19'	-	-	7.67 (m, 1H)	107.12
11'-NH-	-	-	10.28 (s, 1H)	-
1''	-	-	4.47 (t, J = 5.3 Hz, 2H)	49.25
2''	-	-	3.80 (t, J = 5.3 Hz, 2H)	68.76
3''	-	-	4.28 (m, 2H)	59.17
4''	-	-	4.11 (m, 2H)	61.01
5''	-	-	3.48 (m, 2H)	69.56
6''	-	-	3.52 (m, 2H)	69.68
7''	-	-	3.47 (m, 2H)	69.62
8''	-	-	3.38 (t, J = 6.0 Hz, 2H)	69.14
9''	-	-	3.17 (q, J = 5.8 Hz, 2H)	38.42
9''-NH-CO-	-	-	7.8 (s, 1H)	-
10''	-	-	-	172.2
11''	-	-	2.05 (t, J = 7.5 Hz, 2H)	35.08
12''	-	-	1.48 (m, 2H)	25.27
13''	-	-	1.28 (m, 2H)	28.02
14''	-	-	1.43/1.59 (m, 2H)	28.17
15''	-	-	3.08 (m, 2H)	55.39
16''	-	-	4.12 (m, 2H)	61.16
17''	-	-	4.29 (dd, J = 7.7, 5.2 Hz, 1H)	59.25
17''-NH-19''	-	-	6.34 (s, 1H)	-
18''	-	-	-	162.99
18''-NH-19''	-	-	6.40 (t, J = 1.9 Hz, 1H)	-
19''	-	-	4.11 (ddd, J = 7.8, 4.5, 2.0 Hz, 1H)	60.93

**Table S4** Oligonucleotides used in this study, related to STAR methods.

<b>Name</b>	<b>Sequence/Use</b>
<i>hrdB</i> -q FP	CCGTTTCCATCGTTCCGAGA RT-qPCR
<i>hrdB</i> -q RP	ATCTGCCCATCAGCCTTCC RT-qPCR
<i>rox</i> -q FP	GCAGATCCAGCTGATGTCGA RT-qPCR
<i>rox</i> -q RP	TCAGGTACCGGTTGACGTTT RT-qPCR
<i>helR</i> -q FP	GACAGCTCGGTGATCGTCTC RT-qPCR
<i>helR</i> -q RP	ACCAGTGC GTTCGACCTTC RT-qPCR
<i>helR</i> FP	AACCCATGCATATGGACCAGTTGGACGGCACC <i>helR</i> amplification and cloning; <i>NdeI</i> site underlined
<i>helR</i> RP	CCACAAAGCTTTCAGGAGCTGGTGAGGATCACGAG <i>helR</i> amplification and cloning ; <i>HindIII</i> site underlined
<i>helR</i> -FLAG	CCACAAAGCTTTCACTTGTGTCATCGTCCTTGTAGTCGGAGCTGG TGAGGATCACGAG Construction of C-terminal FLAG-tagged HelR; <i>HindIII</i> site underlined
<i>helR</i> -his <sub>6</sub>	CCACAAAGCTTTC AATGATGGTGATGGTGGTGGGAGCTGGTGAGGA TCACGAG Construction of C-terminal hexahistidine tagged HelR; <i>HindIII</i> site underlined
pIJseq FP	ACGTCCATGCGAGTGTCC Sanger sequencing of pIJ10257 constructs
pIJseq RP	CTCCGCTCATGAGAACCCTA Sanger sequencing of pIJ10257 constructs
HelRint	TGACTACTGGGACGAGGTGTTC Sanger sequencing of <i>helR</i> constructs (internal primer)
pGUSseq	AAGCTTGCTCAATCAATCACC Sanger sequencing of pGUS constructs
P <sub>rox</sub> FP	ACTGTCTAGACCGCCAGCATGCCCTACCAAC Amplifying P <sub>rox</sub> ; <i>XbaI</i> site in bold
P <sub>rox</sub> RP	TTAGGGTACCGAGAACC GCCCGTTTCCGC Amplifying P <sub>rox</sub> ; <i>KpnI</i> site in bold
<i>hrdB</i> FP	AACCCATGCATATGTCGGCCAGCACATCCCG Amplifying <i>hrdB</i> ; <i>NdeI</i> site in bold
<i>hrdB</i> RP	CCACAAAGCTTTCAGTCGAGGTAGTCGCGCAG Amplifying <i>hrdB</i> ; <i>HindIII</i> site in bold
P <sub>AP3</sub> FP	AAGCTTATCTATGGATGACC Amplification of P <sub>AP3</sub> for in-vitro transcription
P <sub>AP3</sub> RP	AAGCTTTGTA CTTCGCCCTT Amplification of P <sub>AP3</sub> for in-vitro transcription
<i>helR</i> D533A FP1	GGTAGGATCGTCTAGAACAGGAGGCC Generation of <i>helR</i> with a D533A substitution
<i>helR</i> D533A RP1	AGCTCCTGCGCCTCGGCCACGAC Generation of <i>helR</i> with a D533A substitution
<i>helR</i> D533A FP2	TCGTGGCCGAGGCGCAGGAGCTG Generation of <i>helR</i> with a D533A substitution

<i>helR</i> D533A RP2	<b>CCGGCGTAACAGATGAGGGCAAGC</b> Generation of <i>helR</i> with a D533A substitution
<i>helR</i> K.O. FP	<b>TGCAGAAATGAGACTCACCTCGTGCCTGTGCTTTCGTGCGATTCCGG GGATCCGTCGACC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>helR</i> disruption
<i>helR</i> K.O. RP	<b>CGCCCGGCCCCGGGCGCGGTGCGGCACCCACCCACCTGTAGG CTGGAGCTGCTTC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>helR</i> disruption
<i>helR</i> cPCR FP	<b>GGAAAGGCAAGAGGGTCCAA</b> Amplification of <i>helR</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>helR</i> cPCR RP	<b>CGCTCGTGTGCTGAGTCAT</b> Amplification of <i>helR</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>rox</i> KO FP	<b>CAGCCGTACCAATTGACCTGCGGAAACGGGGCGGTTCTCATTCCGG GGATCCGTCGACC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>rox</i> disruption
<i>rox</i> KO RP	<b>CGGCCGTGGCGGCGCCGAACCACGCGGGCATCCGGGTGATGTAGG CTGGAGCTGCTTC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>rox</i> disruption
<i>rox</i> cPCR FP	<b>GAATCCACAGCCGTACCAATT</b> Amplification of <i>rox</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>rox</i> cPCR RP	<b>GCTCTTCTGCTGCTGTGCG</b> Amplification of <i>rox</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>SVEN_5092</i> KO FP	<b>GAAGTTCACTAAACGCACTGTATCGCGCCGGGAGTCGAAATTCCGG GGATCCGTCGACC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>SVEN_5092</i> disruption
<i>SVEN_5092</i> KO RP	<b>GGCGCACGGCTCTTACCTGCGGGCTGCGGAGCGGATGATGTAGGC TGGAGCTGCTTC</b> Amplification of [ <i>aac(3)IV – oriT</i> ] cassette with homologous regions for <i>SVEN_5092</i> disruption
<i>SVEN_5092</i> cPCR FP	<b>GGAAAGGCAAGAGGGTCCAA</b> Amplification of <i>SVEN_5092</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>SVEN_5092</i> cPCR RP	<b>CGCTCGTGTGCTGAGTCAT</b> Amplification of <i>SVEN_5092</i> region in <i>S. venezuelae</i> genome to confirm disruption
<i>bla</i> FP	<b>CCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTA</b> Amplification of [ <i>hygB – oriT</i> ] cassette with homologous regions for targeting supercos-1 <i>bla</i> gene
<i>bla</i> RP	<b>AATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACC</b> Amplification of [ <i>hygB – oriT</i> ] cassette with homologous regions for targeting supercos-1 <i>bla</i> gene

## **CHAPTER 5**

### **Repurposing inducible rifamycin resistance for the high throughput discovery of RNA polymerase inhibitors**

## **PREFACE**

The work presented in Chapter 5 is a manuscript in preparation for submission.

Surette, M.D., Gong, M.W., Walsh, J.R., Koteva, K., and Wright, G.D.

MDS designed and performed experiments, carried out high-throughput screens, purified and characterized active compounds, wrote code, performed analysis, made figures, and edited and wrote the manuscript. MWG carried out high-throughput screens, purified and characterized active compounds. JRW purified and characterized active compounds. KK performed NMR analysis of Chaxamycin D. GDW supervised the work in addition to writing and editing the manuscript.

## ABSTRACT

The global rise in antibiotic-resistant bacteria has coincided with a failure to develop new antibiotics, precipitating the era of antibiotic resistance. We must discover new antibiotics that are not susceptible to existing mechanisms of resistance to avert this crisis. A single RNA polymerase (RNAP) enzyme performs all transcription in prokaryotes, and this fact is currently exploited by two classes of RNAP inhibitors on the market, rifamycins, and fidaxomicin. These factors make RNAP an attractive target for antimicrobial drug development. Many environmental bacteria produce specific rifamycin resistance enzymes in the presence of these antibiotics. Here, we hijack this inducible system and develop a cell-based high-throughput screen for RNAP inhibitors capable of detecting sub-nanogram quantities of rifampin. We used this to mine a sizeable microbial extract library and combined phenotypic dereplication and whole genome sequencing to prioritize compounds. Using the phenotypic approach, we found both common and rare rifamycins such as rifamycin SV and chaxamycin D. More importantly, this approach also led to the first instance of a rifamycin that can evade rifamycin phosphotransferase (Rph) resistance enzymes. Additional work is required to determine the structure of this compound and the molecular basis of Rph-evasion. Next, sequencing revealed the presence of an unusual rifamycin biosynthetic gene cluster which we hypothesize produces halogenated rifamycins. We describe this cluster's successful capture, activation, and heterologous expression in *Streptomyces coelicolor*. This work will enable the subsequent characterization of this new rifamycin family.

## INTRODUCTION

The emergence and proliferation of antibiotic-resistant bacteria pose a significant threat to public health in the 21<sup>st</sup> century and beyond. These pathogens are no longer susceptible to once-reliable therapies, and their incidence is increasing globally<sup>1</sup>. A report commissioned in the United Kingdom estimated that by 2050 resistance will claim roughly 10 million lives per year, more than cancer (8.2 million)<sup>2</sup>. Addressing this crisis will require more judicious use of these compounds, better global living standards and sanitation, and of course, new antibiotics to replace the agents which are failing<sup>2-4</sup>. In the fifty years from 1929 to 1979, almost twenty new classes of antibiotics entered clinical use. The next forty-four years have produced only four, and almost all major pharmaceutical companies have abandoned their antimicrobials programs<sup>5,6</sup>. This discovery void has played a role in precipitating the antibiotic resistance crisis. The last ~5-10 years have seen investments from governments and nonprofit organizations aimed at rejuvenating the preclinical and clinical antibiotic pipeline but the antibiotics currently in development are insufficient to avert the coming crisis; we urgently require more antibiotics active against resistant strains<sup>7,8</sup>.

Microbial metabolites are the source of most antibiotics used today. The direct screening of bacteria and bacterial products for the ability to inhibit the growth of pathogens fueled the golden age of antibiotic discovery from the 1930s to the 1970s. This approach was remarkably successful but became increasingly challenging due to the frequent re-isolation of known compounds, which are time-consuming to de-replicate, and the slowing pace of novel scaffold discovery<sup>6,9,10</sup>. Ultimately, microbes as a source of

antibiotics have been abandoned by major pharmaceutical companies because discovering new compounds has become too difficult.

More recently, the advent of genomics and synthetic biology has led to a renaissance in microbes as sources of antibiotics. Classic antibiotic-producing genera such as *Streptomyces*, which were once considered exhausted, were revealed through whole genome sequencing to contain the genetic potential to produce far more compounds than we can isolate in laboratory fermentations<sup>11,12</sup>. Synthetic biology has enabled genetic manipulation of these biosynthetic gene clusters and their producers, allowing us to access this cryptic chemical diversity<sup>13-15</sup>. Another promising strategy is so-called high content screening. Trying to find antibiotics by directly screening for growth inhibition from microbial extracts leads to the rediscovery of known nuisance compounds, like streptomycin or streptothricin. By instead looking for specific phenotypes, such as the activation of cellular stress response pathways, one can find molecules that have specific activities, interact with specific cellular targets, or are present in sub-inhibitory quantities<sup>16-18</sup>. An excellent example of this approach is the panel of cell-based mechanism of action reporters described in Wex *et al.* 2021. They use transcriptional fusions of antibiotic-stimulated promoters with a reporter gene to reveal the cellular target during high-throughput screening<sup>19</sup>. With specific antibacterial targets in mind, this approach can allow for targeted interrogation of large natural product libraries with increased sensitivity and fewer nuisance compounds.

Antibiotics inhibit the growth of bacteria because they interfere with essential processes; however, not all essential genes are good antibiotic targets<sup>20,21</sup>. In prokaryotes,

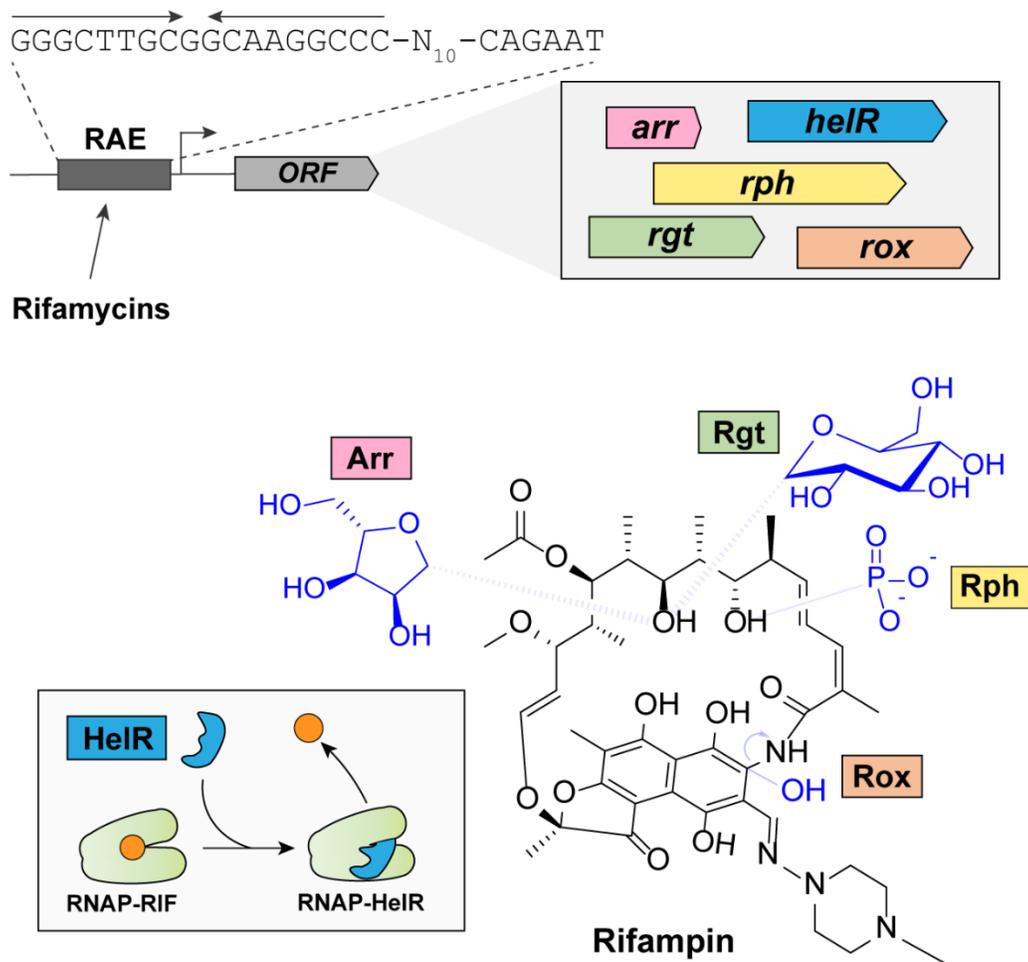
a single DNA-dependent RNA polymerase (RNAP) catalyzes the synthesis of all cellular RNAs, making it an excellent antimicrobial target. Indeed, a wealth of microbial metabolites target this enzyme through multiple distinct modes of action<sup>22,23</sup>. The FDA has approved two classes of antimicrobials targeting RNAP lipiarmycins (fidaxomicin) and rifamycins, both derived from bacteria. Fidaxomicin blocks the movement of mobile domains in RNAP required for melting DNA and initiating transcription. This compound is a narrow-spectrum agent for the treatment of *Clostridioides difficile*<sup>24,25</sup>. Rifamycins, such as rifampin, obstruct the path of the growing transcript inside RNAP, blocking the synthesis of RNAs longer than 2-3nt<sup>26,27</sup>. Members of the rifamycin class are critical for treating *Mycobacterium tuberculosis* and non-tuberculosis mycobacteria (NTM) and cases of severe Gram-positive infections<sup>28</sup>. The discovery of new RNAP inhibitors is a promising avenue of antimicrobial research.

Apart from identifying new chemical scaffolds that can inhibit RNAP, identifying derivatives of known compounds which can overcome resistance mechanisms is equally impactful. Rifamycins suffer from high rates of spontaneous resistance arising from missense mutations in their binding site on the  $\beta$ -subunit (RpoB)<sup>29</sup>. Consequently, these antibiotics are used exclusively in combination therapy to suppress the development of resistance<sup>28</sup>. Notably, specific substitutions appear to be favored. For instance, substitutions in just three residues account for >85% of resistant isolates in *M. tuberculosis*<sup>30</sup>. This homogeneity raises the possibility of designing/discovering rifamycins that can bind to these resistant RpoBs. In 2018, two independent groups found that the rifamycin congener kanglemycins can inhibit RNAP possessing some of the most common

rifampin-resistant substitutions<sup>31,32</sup>. Aside from mutations in the target, which is the only route to resistance for *M. tuberculosis*, many different mechanisms of rifamycin resistance exist in other organisms. For instance, four different classes of rifamycin-inactivating enzymes have been discovered, the rifamycin phosphotransferases (Rph), monooxygenases (Rox), ADP-ribosyltransferases (Arr), and glycosyltransferases (Rgt)<sup>33</sup>. Additionally, we have recently described a fifth mechanism, where rifamycins are displaced from RNAP by the helicase-like protein HelR (Chapter 4)<sup>34</sup>. Some mechanisms are already important clinically, such as Arr, Rox, and HelR enzymes found in fast-growing *Mycobacterium* spp<sup>35</sup>. Others, such as Rph, don't pose an immediate clinical threat but can be found in many rifamycin-susceptible Gram-positive species such as *Listeria*, *Bacillus*, and *Clostridium* spp<sup>36</sup>. Rifamycins, such as TNP-2092, are currently in development to treat *C. difficile* and may select for the expression of this silent resistance determinant<sup>37</sup>.

Curiously, within Actinobacteria, the expression of *rgt*, *rph*, *arr*, *rox*, and *helR* are all regulated at the transcriptional level by the presence of rifamycins (**Figure 1**)<sup>33,36</sup>. Ultimately, the mechanism that controls this rifamycin-mediated induction is unknown but *cis*-regulatory sequences required for induction have been identified, called the rifamycin-associated element (RAE)<sup>36</sup>. More importantly, we have demonstrated this system senses RNAP inhibition by rifamycins (Chapter 2). RNAP inhibitors such as streptolydigin and fidaxomicin do not activate expression, and strains with rifampin-resistant RpoB alleles are defective in induction. Lastly, we show the structurally unrelated antibiotic sorangicin, which binds to the same pocket on the  $\beta$ -subunit as rifamycins, is an equally potent inducer of this system. The RAE not only senses RNAP inhibition but can also discriminate the

specific action of rifamycins and sorangicin from other RNAP inhibitors. In this work, we co-opt this extremely specific antibiotic sensing mechanism to design a cell-based screen for RNAP targeting antibiotics. We present the results of our efforts to screen a small synthetic library of almost 4000 compounds and our in-house collection of >10 000 microbial extracts.



**Figure 1 Inducible rifamycin resistance. A) Rifamycins stimulate the expression of diverse ORFs downstream of the rifamycin-associated element (RAE). B) Resistance mechanisms. HelR removes rifamycins (RIF, orange circle) from RNAP. Products of inactivating enzymes are depicted on rifampin (modifications in blue).**

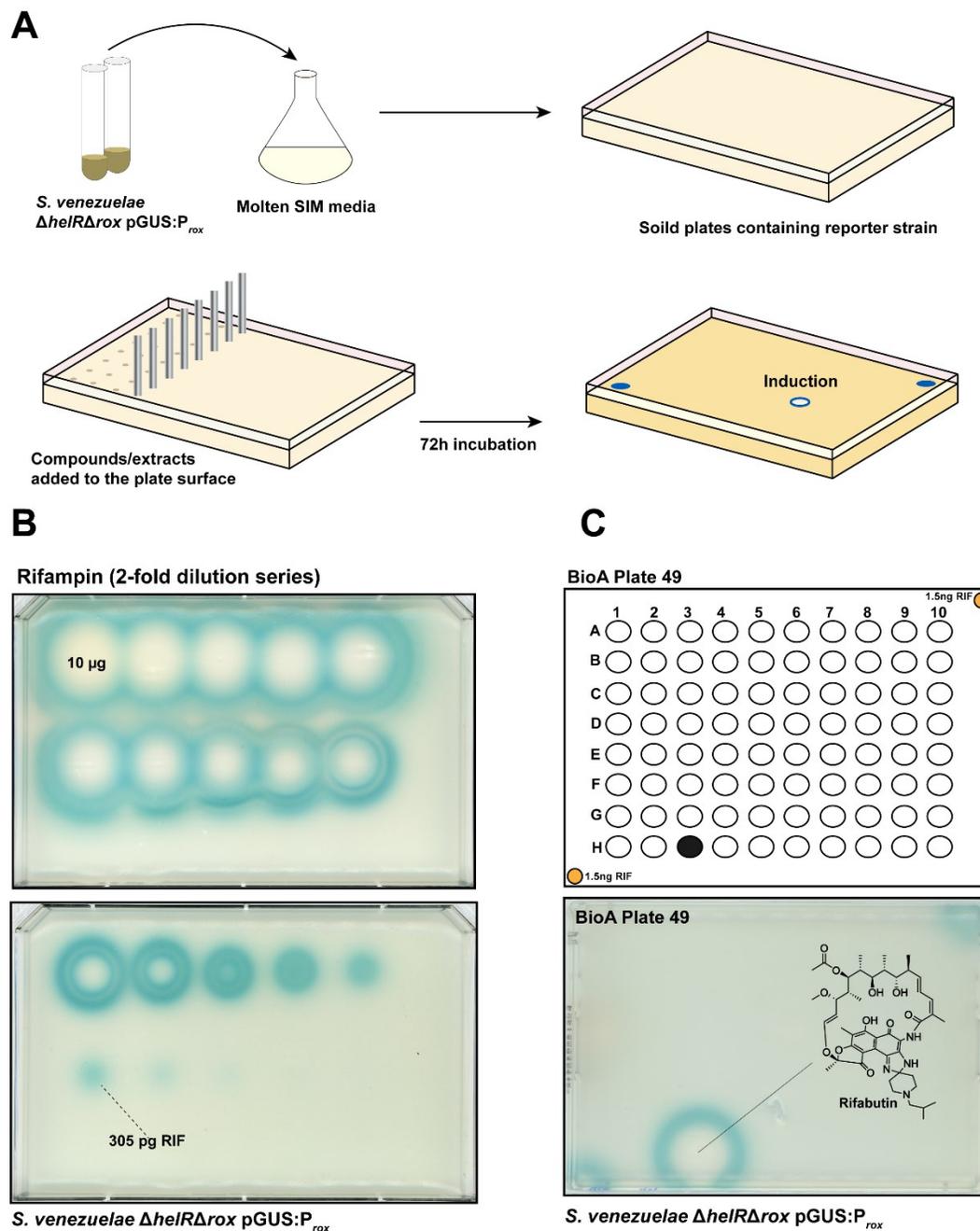
## RESULTS

### A high-throughput screen for RNAP inhibition

To develop a high throughput screen for RNAP inhibitors using the RAE, we used an existing reporter system and the model organism *Streptomyces venezuelae* ATCC 10712. We used the previously validated pGUS: $P_{rox}$  construct, which faithfully recapitulates rifamycin-mediated induction (Chapter 2). This system encodes a transcriptional fusion of  $\beta$ -glucuronidase (GUS) to the RAE-containing intergenic region upstream from the rifamycin monooxygenase *rox*. Expression is detected by adding a colorimetric reagent X-gluc (Alfa Aesar), which GUS metabolizes into a blue precipitate. We previously noted that deleting rifamycin resistance genes from *S. venezuelae*, namely *rox* and *helR*, made cells more susceptible to rifamycins but did not necessarily change their induction pattern (Chapter 4)<sup>34</sup>. Induction still occurred at sub-MIC concentrations, but these MICs were now far lower. In other words, hypersusceptible strains should have a lower detection limit for rifamycins and other RAE-sensitive RNAP inhibitors. Indeed *S. venezuelae*  $\Delta helR\Delta rox$  pGUS: $P_{rox}$  reproducibly detected >100-fold less rifampin than *S. venezuelae* pGUS: $P_{rox}$  (**Supplementary Figure 1**). After unsuccessful attempts to adapt this assay to a liquid format in 96-well plates, we focused on developing an assay using solid media. We made plates impregnated with *S. venezuelae*  $\Delta helR\Delta rox$  pGUS: $P_{rox}$  by adding them directly to molten agar and pouring them into single well plates (**Figure 2A**). We used a Mosquito® liquid handler (SPT LabTech) to dispense compounds/extracts onto the surface of the solidified plate, and induction was scored manually after 72 hours of

incubation. Our detection limit was an impressive ~300 picograms of rifampin using this screening methodology and reporter strain (**Figure 2B**).

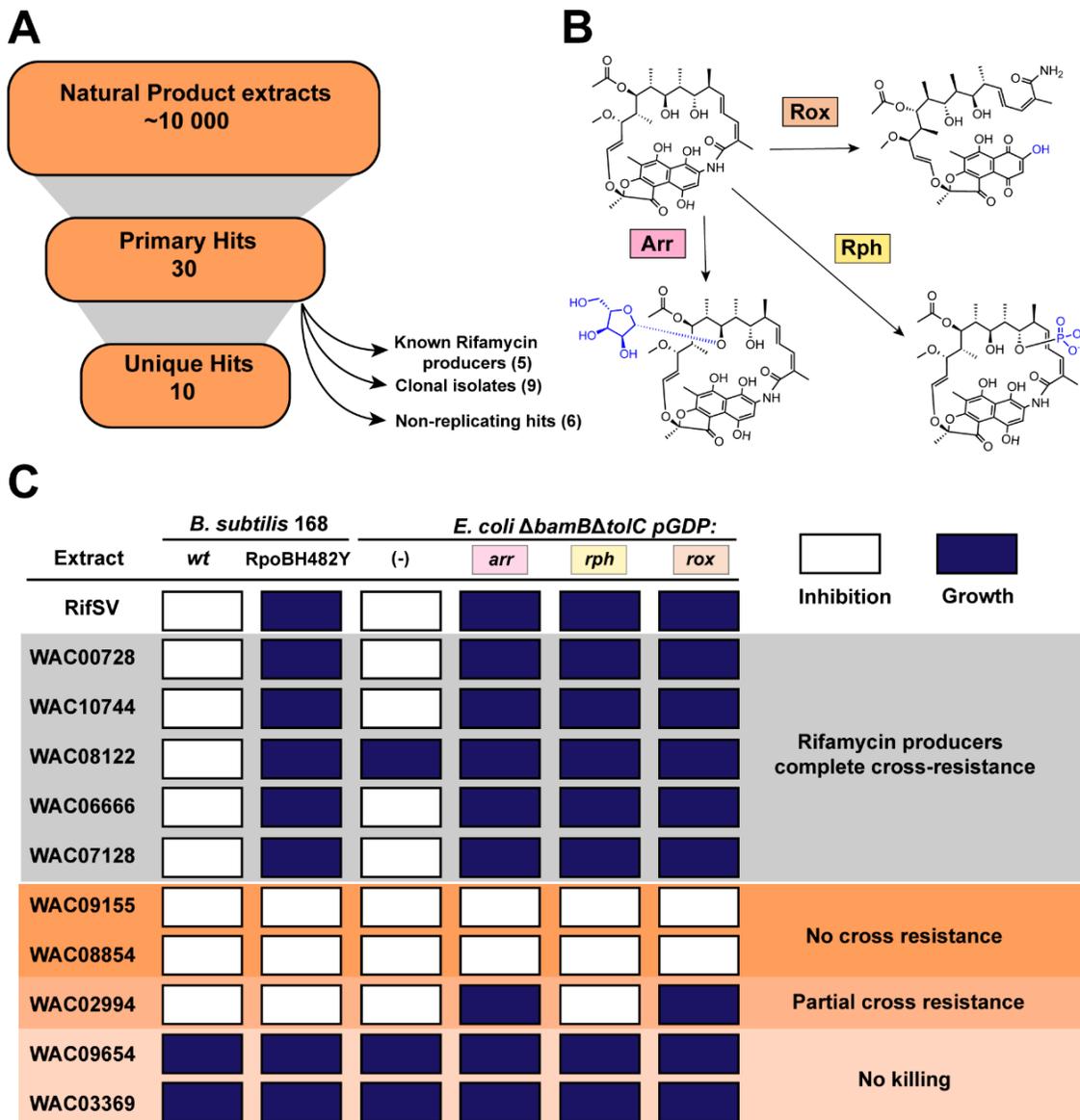
As proof of principle, we began by screening an in-house collection of 3921 curated compounds called the bioactives collection. This collection contains well-known examples of synthetic and natural products with a wide range of reported bioactivities, including but not limited to antibiotic activity (see methods for more details). From the bioactives collection, our assay identified five hits (a 0.13% hit rate), Rifampin, Rifampicin (alternate name for rifampin), Rifaximin, Rifabutin, and Cytochalasin D. Plate 49, which contained rifabutin, is shown as an example in **Figure 2C**. This screen identified all but one rifamycin in the collection, Rifamycin SV, which we have repeatedly demonstrated can induce this reporter (Chapter 2). Inspection of the compound plate revealed that the volume of this specific well was low, likely resulting in a failure to transfer any Rifamycin SV. Our final hit, Cytochalasin D, was intriguing because it is not known to be an inhibitor of prokaryotic RNAP or possess any antibacterial activity<sup>38,39</sup>. However, we ultimately determined this was a false positive, possibly caused by cross-contamination with rifampicin (located in the adjacent well on the compound plate). We could not replicate the induction of the RAE using a new, commercial stock of Cytochalasin D. It also did not affect *S. venezuelae* RNAP during *in vitro* transcription, even at concentrations >300  $\mu$ M (**Supplementary Figure 1**). Overall, our pilot screen of the bioactives library confirmed the performance and validity of our assay. It also confirmed previous data that induction of the RAE is exceptionally specific to rifamycins and sorangicin, as the dozens of antibiotics in this collection all failed to induce our reporter.



**Figure 2 A high throughput screen for RAE activation. A) Solid media screening procedure.** Production of a blue precipitate indicates the induction of our reporter and RNAP inhibition. **B) Limit of detection.** The screening procedure from A) was used to assay a two-fold dilution series of rifampin and determine the smallest quantity we could detect (305 pg). **C) Sample screening plate from the bioactives collection.** Plate layout and assay results for plate 49, containing one of our hits, rifabutin.

### Natural Product Library Screen

Confident in the performance of our assay, we moved on to screening our in-house natural product library (NPL) and associated pre-fractionated library (PFL). The NPL consists of concentrated methanolic extracts from microbes isolated from diverse soils from all over the world. Actinomycetes constitute most of the collection, but it also includes other bacterial taxa and some fungi. The PFL contains methanolic extracts which have undergone crude reverse-phase fractionation. In total, we screened extracts derived from ~10 000 microbes and obtained 30 hits for a rate of ~0.3% (**Figure 3A**). Among these 30 were all four previously characterized rifamycin producers. Another strain, WAC08802, wasn't known to be a rifamycin producer but had been sequenced previously, enabling us to determine that it encoded a rifamycin biosynthetic gene cluster (BGC). We prepared new extracts from the remaining strains, and after removing clonal isolates and strains whose fresh extracts did not activate the RAE, ten unique hits remained (**Supplementary Table 1**). Next, we assayed all ten crude extracts against four different mechanisms of rifamycin resistance to quickly identify rifamycin producers from cross-resistance patterns. We used *Bacillus subtilis* 168 with a common RpoB substitution (H482Y *B. subtilis* numbering) and three strains of rifamycin-sensitive (hyperpermeable) *Escherichia coli*  $\Delta$ *bamB* $\Delta$ *tolC* (*E. coli*  $\Delta\Delta$ ) expressing the inactivating enzymes Rox, Arr, and Rph, which all modify different moieties on the rifamycin scaffold (**Figure 3B**)<sup>33</sup>. We opted for redundancy in the inactivating enzyme panel because rifamycins capable of evading one or more of these enzymes could be useful and likely have a new/unusual structure. Of the ten extracts, four inhibited the growth of wildtype *B. subtilis* and *E. coli*  $\Delta\Delta$  but were



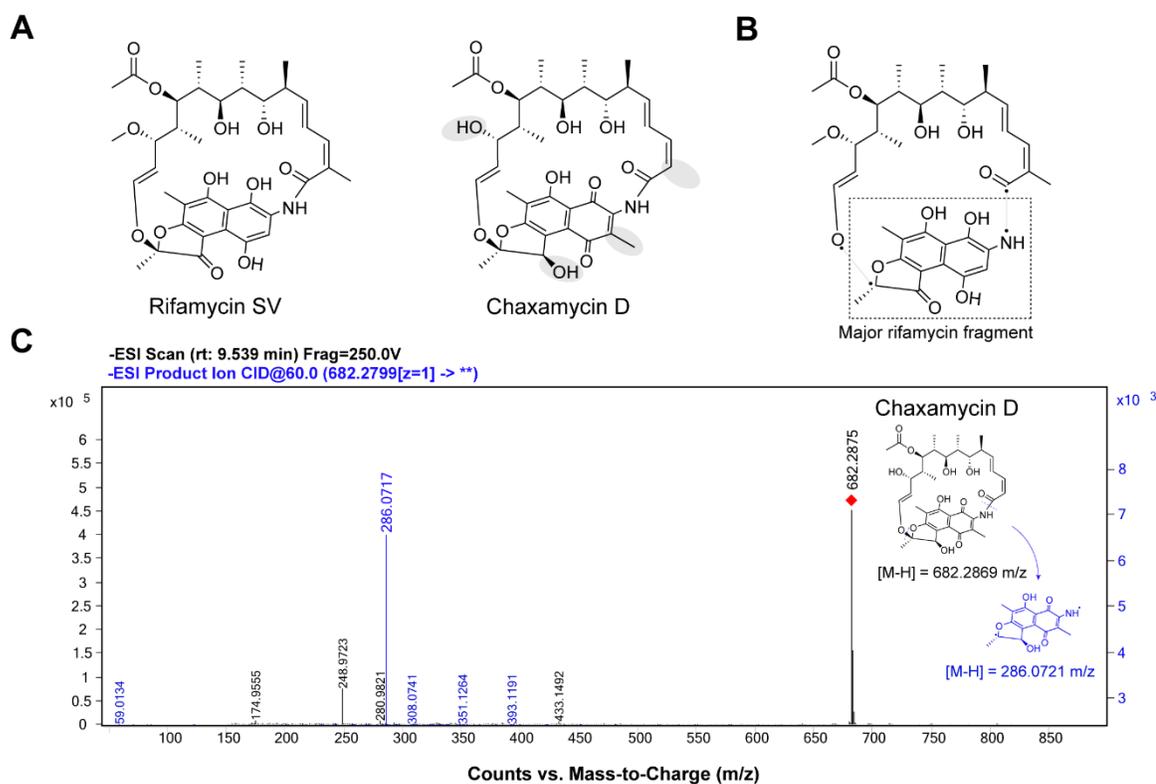
**Figure 3 Natural product library screen and rifamycin dereplication. A) Screening results B) Inactivating enzymes used in dereplication. C) Rifamycin dereplication and hit prioritization.** Extracts were screened for growth inhibition against a panel of bacteria harboring various rifamycin resistance mechanisms and isogenic strains. RpoB numbering is according to the *B. subtilis* enzyme. We used previously validated pGDP1 and pGDP3 constructs of *rox* and *arr/rph*, respectively, for susceptibility testing. pGDP1:*rox* pGDP1 and 3 differ only in the selection cassette. The *E. coli*  $\Delta\Delta$  empty vector control contained pGDP3(-).

inactive against strains expressing all four mechanisms of rifamycin resistance. Accordingly, we designated these rifamycin producers (**Figure 3C**). Crude extract from WAC08122 didn't inhibit *E. coli*  $\Delta\Delta$  and had only modest activity against *B. subtilis*. Still, we included it in the rifamycin producers because it had no activity against the RpoBH482Y strain. More interesting to us were the extracts that showed no cross-resistance, such as WAC09155 and WAC08854, or partial cross-resistance, like WAC02994. These represent the most likely candidates for novel chemical scaffolds as they can evade rifamycin-specific resistance mechanisms. Lastly, fresh extracts prepared from two hits, WAC09654 and WAC03369, gave only very faint activation of our reporter and, unsurprisingly, could not inhibit any of our indicator strains (**Figure 3C**). Only four rifamycin producers have been found in our natural product library in >10 years of screening. By repurposing the RAE for drug discovery, we've more than doubled that number from this screen alone.

### **WAC09155**

We immediately began activity-guided purification of WAC09155 by following induction and killing of *S. venezuelae*  $\Delta helR\Delta rox$  pGUS: $P_{rox}$ . Although none of the rifamycin resistance mechanisms appeared to confer resistance to WAC09155 extracts, we ultimately isolated a rifamycin congener, chaxamycin D, as the active compound. The exact mass determined by high-resolution mass spectrometry (HRMS) and fragmentation analysis supports this assignment (**Figure 4**). Unfortunately, our yield of chaxamycin D (0.8mg) was insufficient for a complete structural assignment with NMR. In support of our HRMS data, whole genome sequencing revealed a biosynthetic gene cluster with high

identity to the chaxamycin cluster from *Streptomyces leeuweenhokii* encoded by WAC09155 (**Supplementary Figure 3**). Chaxamycin D differs from the closely related rifamycin SV by differences in methylation. They lack the typical O-methylation at C27, and the olefin methyl derived from methyl malonyl CoA used in the final module of the PKS (instead of accepting malonyl CoA), and they are methylated at the C3 position on the naphthoquinone core. Lastly, they are unique among ansamycins in the presence of an alcohol instead of a ketone on C11 (**Figure 4A**). All of the above properties are relatively minor modifications of the rifamycin scaffold. Accordingly, chaxamycin D could not evade



**Figure 4 A) Chaxamycin D.** Structural comparison of Chaxamycin D and the prototypical rifamycin SV. **B) Fragmentation of rifamycin antibiotics.** **C) High-resolution mass spectroscopy fragmentation analysis of WAC09155 metabolite.** Chaxamycin D calculated = [M-H]<sup>-</sup> 682.2869 m/z, observed [M-H]<sup>-</sup> 682.2875 m/z (error = 0.88 ppm). Fragmentation is consistent with the mass of the major rifamycin fragment + a methyl group. Calculated [M-H]<sup>-</sup> = 286.0721 m/z, observed [M-H]<sup>-</sup> = 286.0717 m/z (error = 1.4 ppm).

**Table 1 – Antimicrobial Susceptibility testing of chaxamycin D**

	Minimal Inhibitory Concentration*	
	Chaxamycin D	Rifamycin SV
<b><i>E. coli</i> BW25113</b>		
Wildtype	2-4	4-8
$\Delta bamB\Delta tolC$	0.125	0.125
$\Delta bamB\Delta tolC$ pGDP3: <i>arr</i>	>32	>32
$\Delta bamB\Delta tolC$ pGDP3: <i>rph</i>	>32	>32
$\Delta bamB\Delta tolC$ pGDP1: <i>rox</i>	>32	>32
<b><i>B. subtilis</i> 168</b>		
Wildtype	0.125	0.03-0.06
RpoB H482Y	>32	>32

\*Results from two independent tests, all MICs expressed in  $\mu\text{g/mL}$

any rifamycin resistance mechanisms when tested *in vitro* (**Table 1**); its antibiotic activity is comparable to rifamycin SV. It seemed likely to us that the inhibition of our rifamycin-resistant indicator strains must be due to the presence of a second antibiotic. However, during activity-guided purification, the only antibiotic activity we observed was also associated with the induction of our reporter and led us to chaxamycin D.

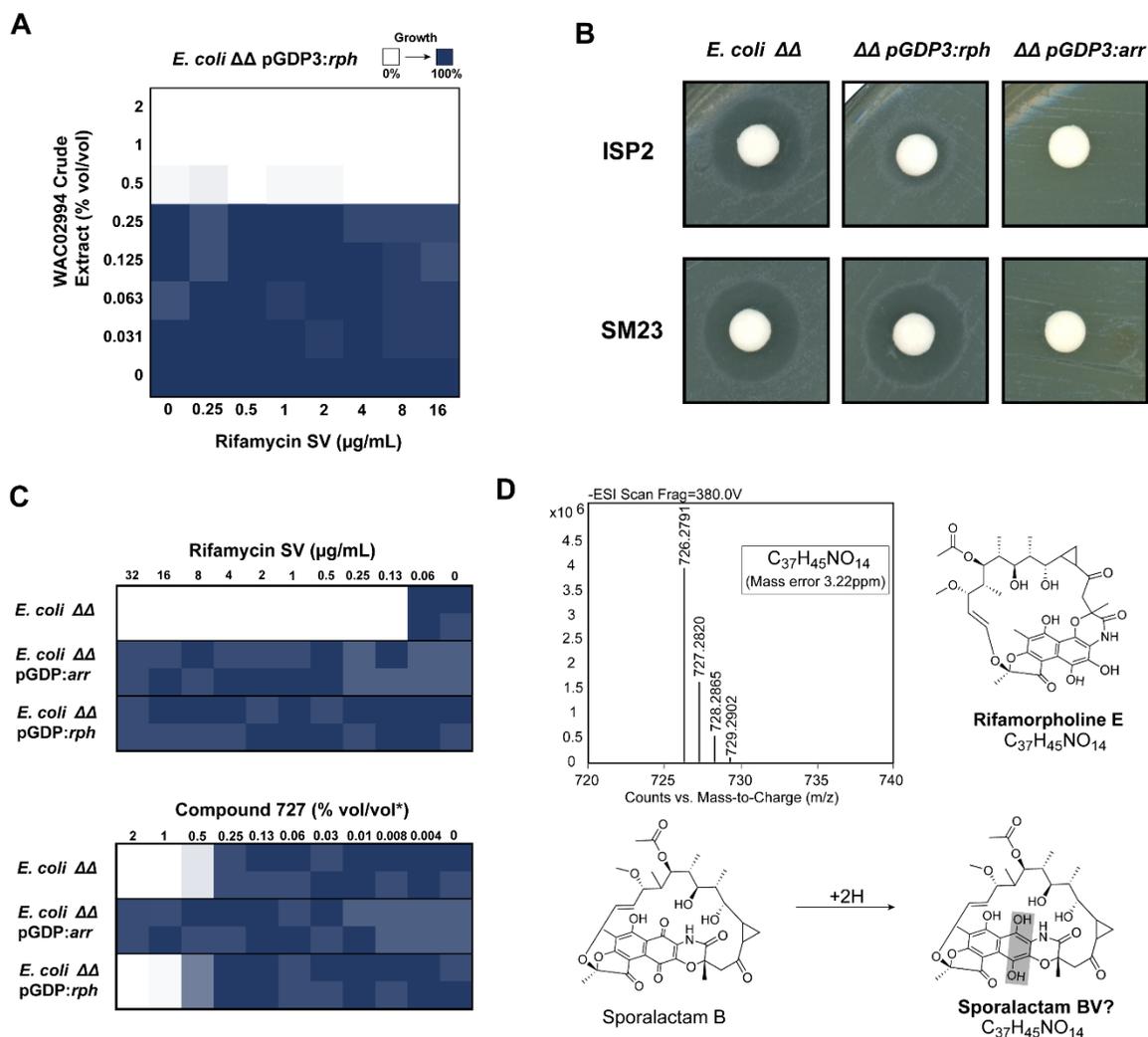
We analyzed our crude extracts by LC-MS coupled with compound dereplication by the Global Natural Products Social Molecular Networking platform (GNPS) to reconcile our initial bioactivity results (**Supplementary Table 1**). This analysis revealed the presence of chloramphenicol in our WAC09155 crude extracts, which we believe is responsible for the blanket inhibition of all strains in **Figure 3**. subsequent whole genome sequencing of WAC09155 confirmed the presence of a chloramphenicol BGC (**Supplementary Table 1**). Co-incidentally, *S. venezuelae* is also a chloramphenicol producer and was, therefore, unaffected by this second antibiotic during our activity-guided purification. Otherwise, the presence of two distinct activities would have been

immediately apparent. Despite originating from a different soil sample, GNPS also detected chloramphenicol in WAC08854, and we found the corresponding ion of chaxamycin D indicating that these isolates are probably close relatives. Chaxamycin production has only been described once in the literature, from *Streptomyces leeuwenhoekii*, isolated from the Atacama Desert in Chile<sup>40</sup>. We found this rare metabolite in multiple strains in our collection using a reporter-based approach. While obtaining rare and diverse microbes has historically been a successful way to find new natural products like chaxamycin D, we demonstrate that targeted screening can unearth rare chemical matter.

#### **WAC02994**

Crude extract from WAC02994 showed a difficult-to-rationalize pattern of susceptibility across resistance mechanisms. Inhibition of *B. subtilis* RpoBH482Y and *rph* expressing *E. coli* suggest a non-rifamycin, whereas the complete protection afforded by *rox* and *arr* implies that inhibition is rifamycin-dependent. GNPS analysis of crude WAC02994 extract revealed the presence of two transcription inhibitors, fidaxomicin and rifamycin SV (**Supplementary Table 1**). Rifamycin SV is susceptible to all inactivating enzymes and cannot engage RpoBH483Y. We, therefore, suspected the unusual resistance phenotypes were fidaxomicin dependent. Testing with a commercial source of fidaxomicin showed that it could inhibit both wildtype and RpoBH482Y *B. subtilis* with a MIC of 1 µg/mL but had no activity against *E. coli* ΔΔ, with or without pGDP3:*rph* (MIC >64 µg/mL). Therefore, the presence of fidaxomicin can explain the activity of WAC02994 extracts against rifampin-resistant RNAP but not against Rph. We first hypothesized that

WAC02994 might make an Rph inhibitor alongside rifamycin SV, which together overcome Rph-expressing bacteria. If this is the case, inhibition of *E. coli*  $\Delta\Delta$  pGDP3:*rph* should depend on rifamycin SV concentration. However, combining crude WAC02994 extract with additional rifamycin SV did not affect the inhibition of *E. coli*  $\Delta\Delta$  pGDP3:*rph*, which instead depended only on the concentration of extract (**Figure 5A**). Based on these data, it seemed most likely to us that WAC02994 instead produces a rifamycin congener which Rph cannot modify. Initial attempts at activity-guided purification of the Rph-evading activity were unsuccessful due to low production levels by WAC02994. After screening >20 different fermentation media, we found SM23, which markedly increased the production of Rph-specific activity (**Figure 5B**). These improved fermentation conditions ultimately enabled activity-guided purification of a small quantity, <0.5mg, of the active molecule using the differential bioactivity against *E. coli*  $\Delta\Delta$  pGDP3:*rph* and pGDP3:*arr*. We've given this compound the temporary name compound 727 after its molecular weight. The small quantity of compound 727 we isolated meant we couldn't confidently determine its potency but were able to confirm its ability to kill *rph* but not *arr* expressing *E. coli* (**Figure 5C**). This compound has UV absorbance at 300 and 400, consistent with a rifamycin, and an observed mass of 727.2869 Da ( $[M-H]^- = 726.27907$  m/z) (**Figure 5D**). We propose the tentative molecular formula  $C_{37}H_{45}NO_{14}$  (3.224ppm mass error) based on reasonable similarity to rifamycins within its size range (35-39C, 1-2N, 10-16O). This molecular formula matches rifamorpholine E and differs from sporolactam B by one saturation unit (**Figure 5D**)<sup>41,42</sup>. Rifamycins can convert through



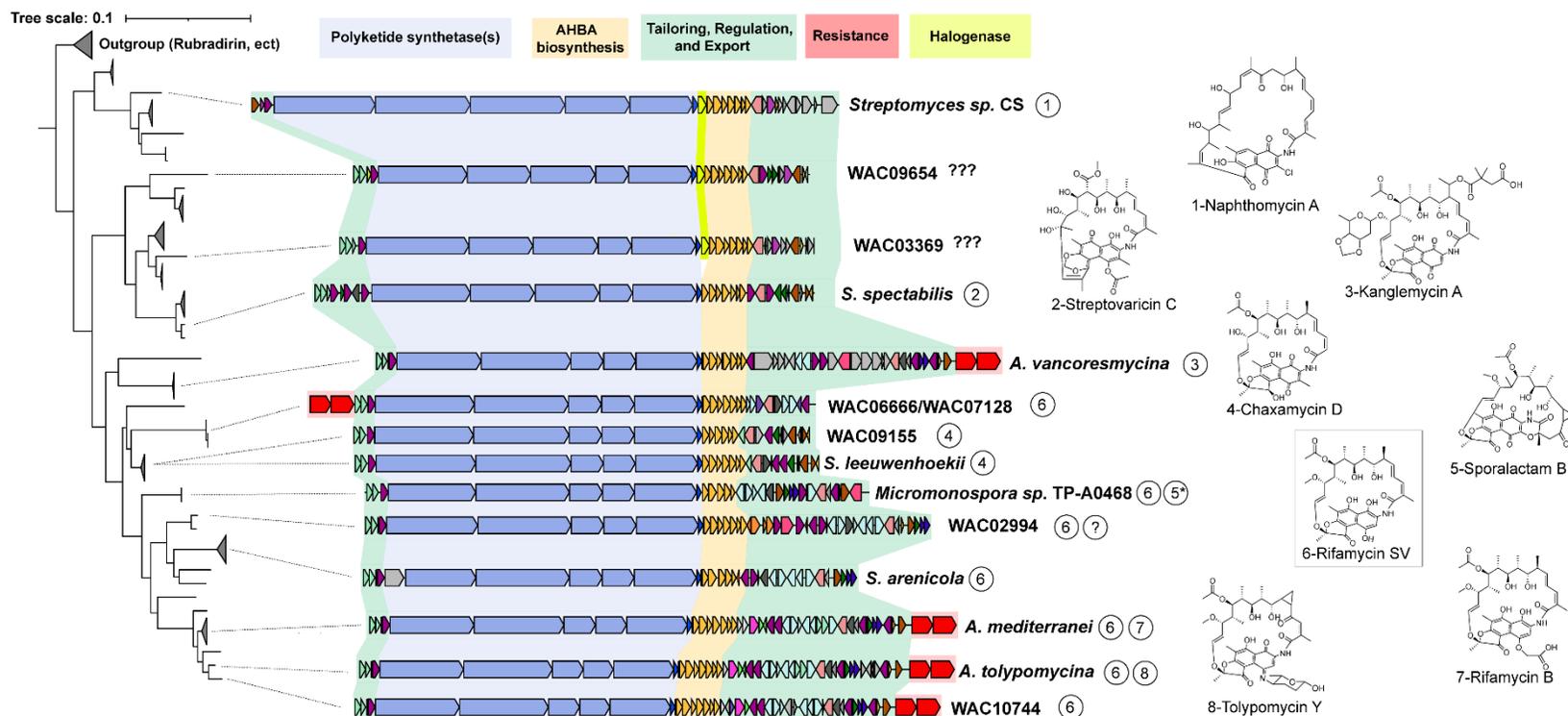
**Figure 5 Purification of an Rph-evasive rifamycin from WAC02994. A) Inhibition of *rph* expressing *E. coli* by WAC02994 extracts is independent of rifamycin SV. A simultaneous concentration gradient of crude extract and rifamycin SV was added to *E. coli*  $\Delta\Delta$  pGDP3:*rph*. Growth inhibition was determined by the OD<sub>600nm</sub> after 16h of growth. B) Increased anti-Rph activity observed in extracts from SM23 media. ISP2, the previous production media, is shown for comparison. C) Compound 727 is responsible for anti-Rph activity. The small quantity of compound 727 we were able to isolate means we cannot be sure about the concentration used in this assay; thus concentration is given in % (vol/vol). D) HRMS analysis of compound 727 and possible structures.**

oxidation or reduction, between a 1,4 dihydroxynaphthlene and 1,4 naphthoquinone unit in their aromatic core, offering a convenient explanation for a 2H increase in mass to arrive at  $C_{37}H_{45}NO_{14}$ . We've termed this hypothetical compound sporolactam BV in homage to the transformation that occurs between rifamycin S and SV (**Figure 5D**). Both rifamycins possess unique 5-6-6-6 fused ring systems and a rearranged *ansa* bridge featuring a cyclopropyl group adjacent to the hydroxyl group targeted by Rph enzymes (**Figure 5D**). Unfortunately, we need more material for an in-depth structural characterization using NMR, so the ultimate structure of compound 727 and the molecular logic underpinning its evasion of Rph remains unknown. Regardless, this is the first report of a rifamycin that can evade Rph enzymes.

### **Whole genome sequencing**

After taking a phenotypic approach to interrogate our hits, we examined their biosynthetic potential at the genetic level. We could isolate high-quality gDNA for all our hits, except for WAC08122 and WAC00728, and sequenced their genomes. We intentionally omitted WAC08854 because of its presumed high similarity to WAC09155. Analysis of the remaining seven genomes using antiSMASH revealed rifamycin-like clusters in all strains. Importantly it also confirmed the production of rifamycin SV and fidaxomicin by WAC02994 and chloramphenicol and chaxamycins by WAC09155 (**Supplementary Table 1, Supplementary Figure 3**). As a first step, we wanted to place the rifamycin-like BGCs isolated here in the larger context of rifamycin biosynthetic diversity. To do this, we extracted rifamycin-like clusters from all Actinobacterial genomes in the RefSeq database and constructed a phylogenetic tree using the 3-amino-5-

hydroxybenzoic acid (AHBA) synthase protein from each cluster in addition to the seven identified here. This approach has been used previously to examine ansamycin biosynthetic diversity because these compounds all use AHBA as the starter unit for their polyketide synthases and, therefore, strictly require this gene<sup>31,43</sup>. This phylogeny performs well at separating distinct compounds based simply on their AHBA synthase sequences (**Figure 6**). We depicted only a single representative cluster from each clade for visual clarity. Just the seven hits from our screen span a considerable amount of phylogenetic and biosynthetic diversity, WAC10744 is closely related to the other *Amycolatopsis* producers (Kanglemycin notwithstanding), WAC02994 clusters with its fellow *Micromonospora* spp. *Actinomadura rifamycinia* harbors the closest BGC to those found in WAC06666 and WAC07128, and the chaxamycin BGCs from *S. leeuwenhoekii* and WAC09155 both cluster together. Strikingly, however, both of our weakly active strains, WAC09654 and WAC03369, fell into clades with no characterized members. While closely related to streptovaricins, all members of these clades contain a halogenase enzyme found in naphthomycin producers (**Figure 6**). Naphthomycins are, as their name suggests, naphthalenic ansamycin antibiotics. They are frequently chlorinated on their naphthoquinone core and feature a longer ansa bridge quite distinct from the rifamycins. Importantly, they lack many essential moieties for interacting with RNAP and likely inhibit growth through a unique and unknown mechanism. These compounds have received comparatively little study, but it is known that chlorination improves their antibacterial activity. Halogenated rifamycins have not been reported to date, so we were intrigued by the structure and bioactivity of these novel molecules.

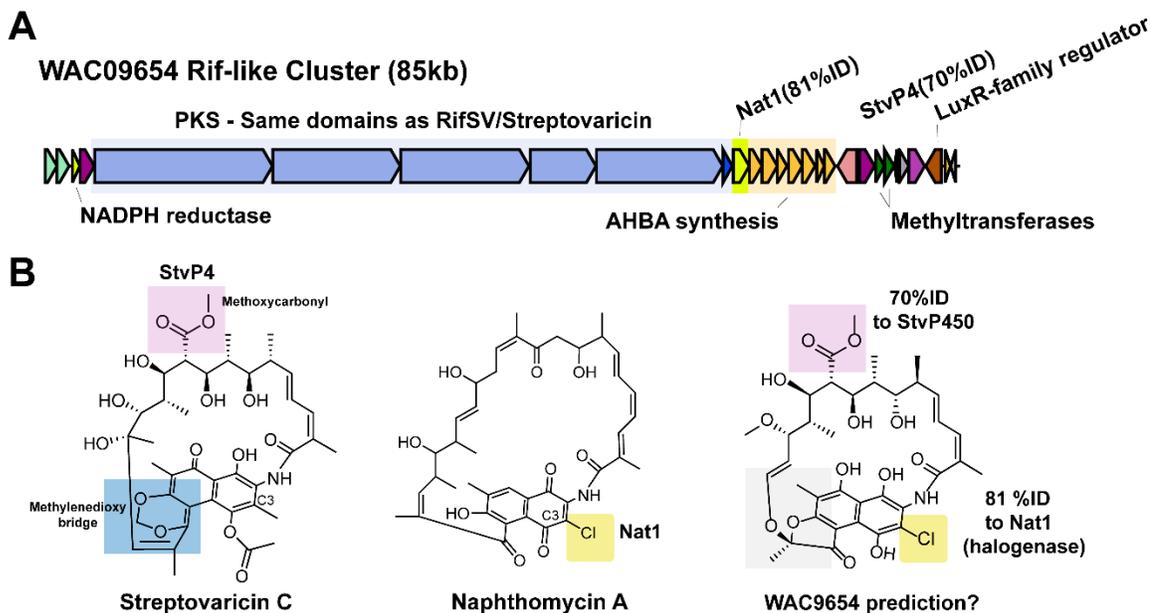


**Figure 6 Structural and genetic diversity of rifamycin-like antibiotics.** Biosynthetic gene clusters for rifamycin-like molecules are placed at the tips of a phylogenetic tree constructed from their conserved AHBA synthase gene. Only select clusters are visualized. Regions of the cluster are color-coded based on function according to the legend at the top. The products of each cluster, where known, are indicated numerically. WAC09654 and WAC03369 belong to clades with no known representatives. \*The sporalactam B producer sequence is not publicly available, but it is stated to be almost identical to the rifamycin cluster from *Micromonospora sp. TP-A0468*<sup>42</sup>. Full species names are *Streptomyces spectabilis*, *Amycolatopsis vancoresmycina*, *Streptomyces leeuwenhoekii*, *Salinispora arenicola*, *Amycolatopsis mediterranei*, and *Amycolatopsis tolypomycina*.

## WAC09654

A close examination of the WAC09654 rifamycin-like BGC revealed an excess of tailoring enzymes compared to WAC03369, so we decided to focus on this cluster. Refer to **Supplementary Figure 4** for detailed functional predictions for each gene in the cluster. First, the halogenase in this cluster is 81% identical at the amino acid level to Nat1, the FMN-dependent halogenase from the naphthomycin cluster, which chlorinates these molecules at the C3 position (**Figure 6**)<sup>44</sup>. The C3 position in the closely related streptovaricins, however, is methylated. WAC09654 encodes two putative methyltransferases, but neither are similar to StvM2 or Cxm24, the putative enzymes responsible for C3 methylation in streptovaricins and chaxamycins respectively, leaving this site conveniently free for chlorination<sup>45,46</sup>. FMN halogenases require a reduced flavin co-factor for activity and therefore have to interact with an NADPH/NADH reductase. Generalist cytoplasmic reductases often fulfill this role, but WAC09654 encodes its own NADPH reductase, similar to one found in the naphthomycin cluster<sup>47</sup>. All of the above factors make it highly likely that WAC09654 produces a C3 halogenated rifamycin. WAC09654 also possesses a homolog to StvP4, which catalyzes the formation of the unique methoxycarbonyl group found in streptovaricins (seemingly in lieu of the acetyl group found on other rifamycins). We presume one of the methyltransferases participates in synthesizing the methoxycarbonyl, but the role of the second is enigmatic. Unlike the streptovaricins, however, WAC09654 lacks the genes required to form the methylenedioxy bridge (StvP2, StvA2, and StvM1), suggesting it may instead form the five-membered ring as seen in the rifamycins<sup>48</sup>. The PKS modules encoded by WAC09654 are consistent with

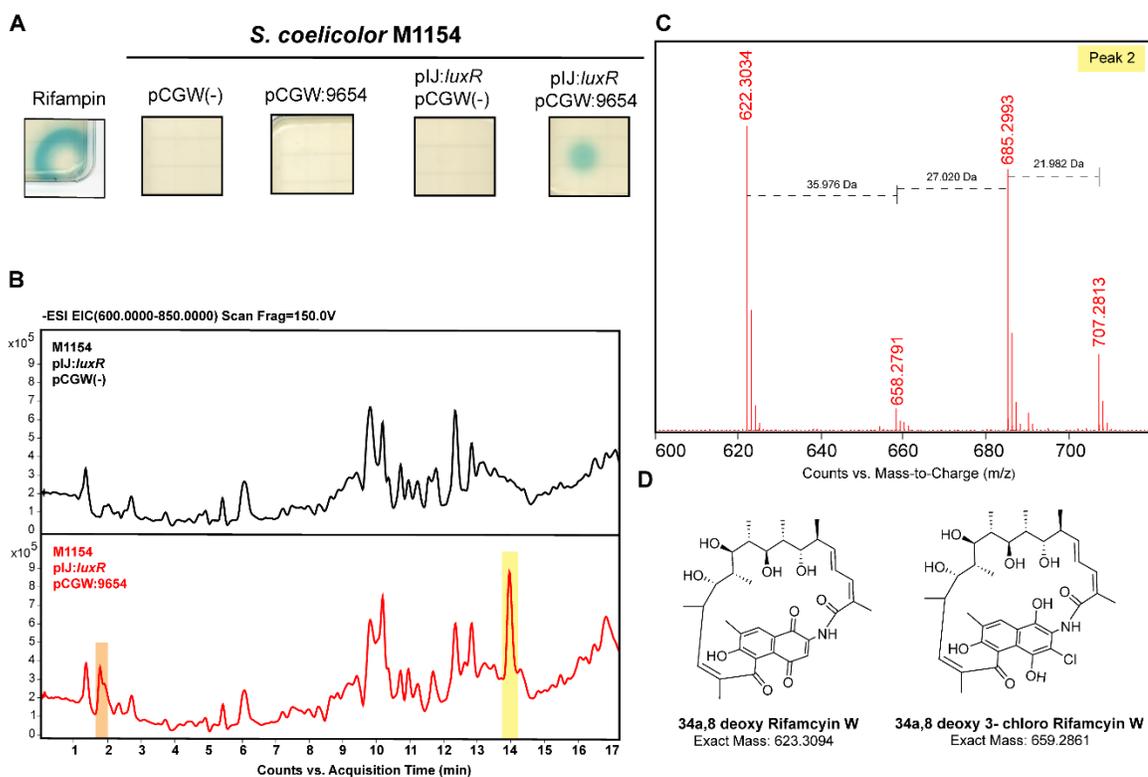
the streptovaricins and rifamycins. Taken together, we propose the structure in **Figure 7** for the putative product of the WAC09654 cluster.



**Figure 7 A) WAC09654 rifamycin-like BGC and B) the predicted product**

Unfortunately, the activation of our RNAP-inhibition reporter by WAC09654 extracts was very weak, and attempts to boost production by different fermentation conditions were unsuccessful. We instead captured the ~85kbp cluster using transformation assisted recombination (**Supplementary Figure 5**) and transferred it by conjugation to *Streptomyces coelicolor* M1154, a common host for heterologous expression of natural products (**Figure 8**)<sup>49</sup>. We confirmed the success of the capture by restriction mapping and subsequently using Illumina Sequencing (**Supplementary Figure 5D and E**). Importantly, M1154 harbors an S433L substitution in its RpoB, associated with increased production of specialized metabolites. For our purposes, it also confers high-level rifamycin resistance and should prevent any toxicity from the WAC09654 cluster.

We performed 50mL fermentations of M1154 and M1154 pCGW:9654, extracted them with ethyl acetate, and dried and resuspended them to be 100-fold concentrated relative to the culture supernatant. When analyzed by HPLC or LC-MS, we couldn't detect any new metabolites produced by M1154 pCGW:9654 (Data not shown). Furthermore, these concentrated extracts failed to activate our RNAP inhibition reporter (**Figure 8A**). Most often, failure of heterologous expression corresponds to a transcriptionally silent cluster<sup>14</sup>. We noted the presence of a predicted LuxR family regulator within the cluster (*ctg1\_662*, **Supplementary Figure 4**). These elements act as positive regulators in other ansamycins<sup>50</sup>. We supplemented M1154 pCGW:9654 with a second copy of its LuxR regulator under the control of the strong constitutive promoter  $P_{ErmE^*}$  (pIJ:*luxR*). This approach was successful as extracts from M1154 pCGW:9654 pIJ:*luxR* activated our reporter (**Figure 8A**). However, the appearance of new metabolites was still not evident by HPLC or LC-MS, and we further optimized the fermentation conditions for maximal production. As other groups have reported for *A. mediterranei*, we found that adding 80 mM KNO<sub>3</sub> led to a marked increase in production<sup>51</sup>. Under these fermentation conditions, we could finally observe the production of two new peaks by LC-MS (**Figure 8B**). The first eluted almost immediately from the column and generated an ion [M-H]<sup>-</sup> of 683.2288 m/z, the second peak was composed of four individual compounds ranging from 622-708 m/z (**Figure 8C**), and all of the novel metabolites fall within the size range of known rifamycins. Of the four compounds identified in the second peak, only one, the smallest, had a mass matching a known rifamycin [M-H]<sup>-</sup> 622.3034 m/z, which we have putatively assigned to 34a,8 deoxy Rifamycin W (**Figure 8D**, mass error = -1.93ppm).



**Figure 8 Heterologous expression of WAC09654. A) Induction of *S. venezuelae*  $\Delta$ helR $\Delta$ rox pGUS:P<sub>rox</sub> by *S. coelicolor* M1154 containing the WAC09654 BGC.** Concentrated ethyl acetate extractions from 50mL fermentations were assayed for induction. Overexpression of the cluster-situated LuxR regulator was necessary to produce RNAP-inhibiting metabolites. **B) Novel metabolites from M1154 pIJ:luxR pGCW:9654.** Ethyl acetate extracts of the heterologous producer and control strain were analyzed by HRMS coupled to a reverse-phase liquid chromatography system. **B)** Depicts the extracted ion chromatogram of 600-850 m/z to highlight the appearance of rifamycins. **C) HRMS of metabolites making up Peak 2.** Mass differences between the various ions are noted in the figure. **D) Structure of putative compounds produced by M1154 pIJ:luxR pCGW:9654.**

This molecule is thought to be a very early post-PKS intermediate in rifamycin biosynthesis<sup>52</sup>. Since WAC09654 appears to have the same PKS modules as rifamycins/streptovaricins, we expect this metabolite to be present, albeit not as a major product (**Supplementary Figure 6**). The next ion, [M-H]-658.2791 m/z, is consistent with

the reduction of the naphthoquinone group and the addition of a chlorine atom, 34a,8 deoxy-3-chloro rifamycin W (**Figure 8D**, mass error = -0.45 ppm). Furthermore, this peak is associated with an isotopic distribution suggestive of chlorine (increased abundance of the [M+2-H]- peak). It must be mentioned that the simultaneous elution of the reduced and oxidized forms of a different structure (which differ by 2 Da) could produce a similar pattern. However, the exact mass is an excellent match for the chlorinated product. The naphthoquinone is a logical location for a reduction, but it could theoretically happen at an alternate site in the molecule. The final ions, [M-H]- 685.2993 and 707.2813 *m/z*, do not correspond to any known rifamycins; we could not deduce candidate structures from the available information. Importantly, neither of these ions (or the one found in Peak 1) has the isotopic distribution of a chlorinated compound (**Figure 8C**). Importantly, we do not know which of these metabolites are inhibiting RNAP. The relative scarcity of halogenated products is surprising, but nevertheless, these do represent novel rifamycin congeners. We have laid the groundwork for future studies to tackle these molecules' purification and structural and biological characterization.

## DISCUSSION

Here we describe and develop an extremely sensitive cell-based assay for mechanism-specific RNAP inhibition by rifamycins by exploiting a widespread antibiotic-sensing pathway associated with resistance. This assay can reliably detect as little as ~300 pg of rifampin (**Figure 2**). From a defined chemical library of >3000 bioactive molecules and a natural product library containing extracts from over 10 000 bacteria (each containing scores of compounds), we identified only natural and synthetic rifamycins and related

analogs, emphasizing the remarkable specificity of the RAE. This screen identified many rifamycins in our library, which have evaded detection through numerous screens on these extracts (**Figure 3**). This approach, therefore, has the potential to act as a cheap and fast dereplication tool for rifamycin producers, as our assay can detect sub-inhibitory amounts (unlike resistance-guided dereplication) and can find novel structures (unlike mass spectrometry). Although we did not uncover any novel RNAP targeting scaffolds, we still believe this is possible with this reporter, as evidenced by its activation by sorangicin.

We took the hits from our natural product library and used phenotypic dereplication with four different mechanisms of rifamycin resistance to prioritize novel and interesting compounds (**Figure 3**). In some respects, this approach was successful; we identified five strains producing a pan-susceptible rifamycin, and confirmed production of rifamycin SV in three of these strains (**Supplementary Table 1**). On the other hand, our phenotypic dereplication strategy failed to flag WAC09155 as a producer of chaxamycin D, a rifamycin congener unable to bypass any rifamycin resistance mechanisms (**Figure 4, Table 1**). Our strategy failed because WAC09155 produced a second antibiotic, chloramphenicol, which we missed during activity-guided purification due to the intrinsic resistance of *S. venezuelae*. Microbial polypharmacy poses a significant hurdle for phenotypic dereplication; we also saw this in WAC02994, which produces both fidaxomicin and rifamycins.

Most importantly, our phenotypic dereplication approach led to the discovery that WAC02994 produces both rifamycin SV and compound 727, a rifamycin that is not susceptible to the rifamycin phosphotransferase Rph. Based on HRMS, we speculate that

it could be a 5-6-6-6 ring system containing rifamycin derivative such as rifamorpholine E or a reduced form of sporalactam B (**Figure 5**). Rifamorpholines were isolated from an insect-associated *Amycolatopsis*, and intriguingly the authors report no rifamycin analogs from this strain, only rifamorpholines, despite the almost identical cluster content and organization to the rifamycin cluster from *Amycolatopsis mediterranei*<sup>41</sup>. In the case of sporalactam B the producer also made rifamycin SV, and like compound 727 in WAC02994, sporalactam B was a relatively minor product. Genome sequencing showed that WAC02994 is a *Micromonospora spp.*, with a typical rifamycin SV BGC (**Supplementary Figure 3, Supplementary Table 3**). Interestingly, the sporalactam B producer belongs to the same genus and, while the corresponding BGC has not been published, it is also described as a relatively unremarkable rifamycin SV cluster, with no major differences from clusters in related species<sup>42</sup>. For these reasons, I speculate that compound 727 is the reduced form of sporalactam B. However, additional work is required to obtain enough compound 727 for comprehensive <sup>1</sup>H and <sup>13</sup>C NMR analysis and complete structural assignment. A structure will also help shed light on the mechanism of Rph evasion, potentially enabling the rational design of rifamycins to evade these enzymes.

Intriguingly, a previous report on the chaxamycins suggested that these compounds don't target RNAP and must have a distinct mode of action because the producer encodes a rifamycin-susceptible polymerase<sup>46</sup>. However, the extreme structural similarity between rifamycin SV and chaxamycin D implies a shared mechanism of action, and its discovery in this screen using the RAE all but confirms it. Upon closer inspection, the chaxamycin producers, *S. leeuwenhoekii* and WAC09155, don't actually possess rifamycin-susceptible

polymerases. These bacteria have a RpoBS456N substitution, found sporadically throughout the actinomycetes and associated with resistance<sup>53</sup>. This substitution confers considerable resistance. *S. coelicolor* (RpoBS456N) has a reported rifampin MIC of ~10 µg/mL<sup>54</sup>, compared to *S. venezuelae* (RpoBS456), which has an MIC 0.5 µg/mL despite two dedicated rifamycin resistance genes, *helR* and *rox*. When these genes are deleted, the MIC falls to 0.063 µg/mL<sup>34</sup>. In fact, the quantity and specific substitutions associated with rifamycin resistance vary significantly across rifamycin producers (**Supplementary Figure 7**). Apart from the chaxamycins, WAC09654, WAC03369, and some *Micromonospora spp.* appear to have only the RpoBS456N substitution suggesting they might be more susceptible to their own antibiotic than the *Amycolatopsis spp.*, which encode several substitutions (**Supplementary Figure 7**). Perhaps this influences the production level of some hosts. Low production titers have been a consistent hurdle during this study. Regardless, all producers of genuine rifamycin-class antibiotics, including the chaxamycins, have, at the very least, the single S456N substitution for self-resistance.

Whole genome sequencing of our hits and comparison to rifamycin-like clusters deposited in public sequence databases revealed the presence of two large clades which appear to encode for C3-chlorinated rifamycins (**Figures 6 and 7**). We successfully captured the cluster from WAC09654 using TAR cloning and activated the cluster in *S. coelicolor* M1154 by overexpressing a pathway-situated LuxR family regulator (**Figure 8 and Supplementary Figure 5**). Preliminary HRMS of WAC09654 BGC-derived metabolites found multiple compounds but only one putatively chlorinated species (**Figure 8**); its mass suggests it is 34a, 8 deoxy 3-chloro rifamycin W. A similar structure to inactive

biosynthetic intermediates of rifamycins (**Supplementary Figure 6**)<sup>52</sup>. Why chlorination would only be present in these theoretically inactive intermediates isn't obvious. Larger rifamycin analogs were produced (683-707 *m/z*), which we *assume* are farther along in the biosynthetic pathway. Until we have purified these compounds, we cannot determine their structures and do not know which one(s) can inhibit RNAP. The naphthomycin producer *Streptomyces sp.* CS makes an excess of non-chlorinated naphthomycin, despite being less active than their chlorinated analogs<sup>44</sup>. It is possible that halogenation of 34a,8-deoxyrifamycin W is inefficient, leading to low-level incorporation in mature rifamycins, which we cannot currently detect (**Supplementary Figure 6**). Future work could include overexpression of the halogenase to address this possibility. Perhaps the halogenated molecule observed in this work has a distinct activity, more similar to the naphthomycins, which also lack an OH at C8. At the same time, maybe the non-halogenated intermediates proceed down the biosynthetic pathway to rifamycins. The heterologous expression strains we created and validated in this study set the stage for investigating these intriguing compounds in future studies.

## METHODS

Table 2 – Strains and isolates used in this work

Strain or Plasmid	Genotype/Use	Source or Reference
<b><i>Streptomyces venezuelae</i></b>		
ATCC 10712		ATCC
pGUS:P <sub>rox</sub>		Chapter 4 <sup>34</sup>
$\Delta$ helR $\Delta$ rox		Chapter 4 <sup>34</sup>
$\Delta$ helR $\Delta$ rox pGUS:P <sub>rox</sub>	High throughput screen for RNAP inhibition	Chapter 4 <sup>34</sup>
<b><i>Streptomyces coelicolor</i></b>		
M1154	Heterologous expression of BGCs	49
M1154 pCGW		This work
M1154 pCGW:9654		This work
M1154 pCGW pIJ:lux		This work
M1154 pCGW:9654 pIJ:lux	Heterologous expression of WAC09654 BGC	This work
<b><i>Escherichia coli</i></b>		
DH5 $\alpha$	General cloning	Invitrogen
ET12567/pUZ8002::bla	Methylation deficient strain used for conjugation into <i>Streptomyces</i>	55
Epi300	Maintenance and copy number control of pCGW	Epicentre
BW25113	K12-derivative susceptibility testing	CGSC 7636
BW25113 $\Delta$ bamB $\Delta$ tolC pGDP3(-)	Hyperpermeable, enhanced susceptibility to rifamycins	9
BW25113 $\Delta$ bamB $\Delta$ tolC pGDP3:arr	Expression of rifamycin ADP-ribosyltransferase	9
BW25113 $\Delta$ bamB $\Delta$ tolC pGDP3:rph	Expression of rifamycin phosphotransferase	9
BW25113 $\Delta$ bamB $\Delta$ tolC pGDP1:rox	Expression of rifamycin monooxygenase	56
<b><i>Saccharomyces cerevisiae</i></b>		
VL6-48N	Highly transformable strain used for TAR	57
<b><i>Bacillus subtilis</i></b>		
ATCC 23857 (Strain 168)	Susceptibility testing	ATCC
RpoBH482Y	Susceptibility testing	58

<b>Plasmids</b>		
pGUS: <i>P<sub>rox</sub></i>	Reporter of RNAP inhibition by rifamycins	Chapter 4 <sup>34</sup>
pIJ10257	Integrative <i>Streptomyces</i> shuttle vector with strong constitutive promoter <i>P<sub>ermE*</sub></i> , Hyg <sup>R</sup>	59
pIJ: <i>luxR</i>	Constitutive expression of 9654 pathway situated <i>luxR</i>	This work
pCGW	TAR cloning vector, Kan <sup>R</sup> , Trp <sup>+</sup>	60
pCGW:9654	Heterologous expression of 9654 BGC	This work
pGDP3: <i>arr</i>	Expression of <i>arr</i> , kan <sup>R</sup>	9
pGDP3: <i>rph</i>	Expression of <i>rph</i> , kan <sup>R</sup>	9
pGDP1: <i>rox</i>	Expression of <i>rox</i> , kan <sup>R</sup>	56

## Chemicals

Unless stated otherwise in the text, all chemicals were purchased from Sigma Aldrich.

## High throughput screen

*S. venezuelae*  $\Delta$ *helR* $\Delta$ *rox* pGUS:*P<sub>rox</sub>* was cultured in 3mL of Tryptic Soy Broth (TSB, BD) + 50 $\mu$ g/mL Apramycin with two 3mm sterile glass beads with shaking (250rpm) at 30°C for 48 hours. The OD<sub>600nm</sub> of these starter cultures was determined after homogenizing the cells by pipetting up and down with a P1000 several times. This culture was added to molten SIM agar (1% soluble starch, 0.03% casein, 0.2% KNO<sub>3</sub>, 0.2% NaCl, 0.2% K<sub>2</sub>HPO<sub>4</sub>, 0.005% MgSO<sub>4</sub> · 7H<sub>2</sub>O, 0.002% CaCO<sub>3</sub>, 0.001% FeSO<sub>4</sub> · 7H<sub>2</sub>O, 1.5% agar, pH 7.2), kept at 45°C in a water bath, to a final concentration of OD<sub>600nm</sub> = 0.1 and X-gluc (Alfa Aesar), and Apramycin (Sigma Aldrich) was added to 80  $\mu$ g/mL and 50  $\mu$ g/mL respectively. On a level surface, 25mL of the inoculated media was added to an OmniTray single well plate

(Nunc™) which was tilted perpendicular to each side of the plate to spread the media and then allowed to dry with the lid open on a level surface for 30 minutes. Plates were kept in sealed plastic bags until their use to avoid shrinking the agar due to moisture loss. A Mosquito® liquid handler (SPT Labtech) was used to transfer 1 µL of compound (Bioactives library) or extracts (NPL/PFL) directly onto the surface of the plate. To achieve success with this protocol, the preparation of level plates with a consistent height of the agar layer was essential. Once compounds had been added to the plates, they were incubated for three days at 30°C, and each plate was scanned with a white and black background to highlight induction (blue pigment) and growth inhibition, respectively. Induction was both obvious and relatively rare, making it feasible to score the plates manually. All other uses of the *S. venezuelae*  $\Delta$ helR $\Delta$ rox pGUS:P<sub>rox</sub> to assess induction was done according to this protocol, except compounds/extracts were added to the plate by hand.

### **Bioactives library**

This in-house library is an amalgamation of four smaller libraries that were cherry-picked to omit duplicate compounds. A subset of the Prestwick Chemical Library®, which contains 539 FDA-approved drugs known to be safe and active in humans. 501 compounds from the BIOMOL2865 Natural Products Library (Enzo Life Sciences), which includes various classes such as peptides, terpenoids, peptolides, flavones, coumarins, alkaloids, macrolides, synthetic derivatives, and more. 1240 compounds from the Library of Pharmacologically Active Compounds international version (Lopac®1280, Sigma Aldrich). This collection is curated to include molecules with activities in eukaryotic cell

signaling pathways and neurological models. Lastly, 1641 compounds from the Spectrum Collection (MicroSource Discovery Systems), which is a mixture of known drugs and drug-like molecules and diverse and unusual natural products. All compounds are maintained in 384 well plates as 20 mM stocks in DMSO and stored at -30°C when not in use.

### ***In-vitro* transcription**

Multiple round runoff *in-vitro* transcription reactions were performed using *S. venezuelae* RNA polymerase and  $\sigma^{\text{HrdB}}$  and a fragment of the constitutive mycobacterial promoter  $P_{AP3}$  exactly as previously documented<sup>34</sup>.

### **Crude extract procedure**

To prepare fresh extracts of hits from our screen, we first grew each strain on Bennett's media (1% potato starch, 0.2% casamino acids, 0.18% yeast extract, 0.2% Czapek mineral mix (10% KCl, 10% MgSO<sub>4</sub> · 7H<sub>2</sub>O, 12% NaNO<sub>3</sub>, 0.2% FeSO<sub>4</sub> · 7H<sub>2</sub>O, 0.2% concentrated HCl), 1.5% agar, pH 6.8), to confirm purity. Next, we grew each strain in 3mL of *Streptomyces* Antibiotic Activity Medium (SAM) (1.5% glucose, 1.5% soya peptone, 0.5% NaCl, 0.1% yeast extract, 0.1% CaCO<sub>3</sub>, 0.25% glycerol, pH 6.80) containing a single sterile glass bead for six days at 30°C and 250rpm. These six-day-old cultures were used to streak a confluent lawn of each strain onto a Bennetts agar plate and incubated for another six days. These agar plates were homogenized, resuspended in 10mL of methanol, and shaken overnight at 4°C in a 50mL Falcon<sup>®</sup> Tube (Corning). This slurry was filtered to remove residual pieces of agar. The liquid was removed using a centrifugal solvent

evaporator (Genevac/SP Scientific), dry material was resuspended in 1.5mL DMSO and vortexed with a single glass bead to break up solid pieces. The extract was incubated for 24 h to solubilize as much material as possible. Insoluble material was removed by centrifugation at 17 000 x g for 15 minutes, and the supernatant (crude extract) was transferred to a fresh tube and stored at -20 until further use. All the extracts in the natural product library were prepared in this manner. The pre-fractionated library was created using crude extracts, prepared as described above, applied to a 15 g C18 Gold reverse phase column (RediSep Rf, Teledyne), and fractionated using a linear gradient of acetonitrile. Individual fractions are dried, resuspended in DMSO, and pooled to give eight fractions per extract.

### **High-resolution mass spectrometry (HRMS) and GNPS analysis**

HRMS were acquired using an Agilent 1290 Ultra Performance Liquid Chromatography (UPLC) and Agilent qTOF 6550 mass detector in series. Data were acquired in positive and negative ion modes, but the negative mode generally had improved sensitivity for rifamycins and related compounds. For routine analysis, we used an Eclipse XBD-C8 column (2.1 x 100mm, 3.5 $\mu$ m) for separation with the following method: 0-1 min, 95% solvent A (water + 0.1% formic acid); 1-17 min, a linear gradient from 5% to 95% solvent B (acetonitrile + 0.1% formic acid); 17-18min a linear gradient from 95% to 5% B, 18-19 min 5%B, at a flow rate of 0.5 ml/minute.

For GNPS and general fragmentation analysis, we ran an AutoMS2 acquisition using the advanced decision engine in Agilent MassHunter Workstation V10.0, with three

fixed collision energies (0, 30, and 60 MeV) and an isolation width of ~4 amu. Thresholds for precursor selection were set at 50000 (absolute) or 0.01% (relative) spectral abundance. The same UPLC, qTOF system, and column were used as described above but with the following method for separation: 0-2 min held at 10% B, 2-17min linear gradient from 10-100% B, 17-19 min held at 100%B, 19-20min linear gradient from 100-10% B. with a flow rate of 0.4mL/min. For GNPS analysis, blanks of an appropriate solvent were run between each sample, and spectra were collected in positive and negative ion modes. Raw data were converted to mzXML files and analyzed on the METABOLOMICS-SNETS-V2 server (UCSF). We performed dereplication using the Library Search function on default settings<sup>61</sup>. Relevant hits detected by GNPS were Rifamycin SV (WAC10744, WAC06666, WAC07128, WAC02994), Chloramphenicol (WAC09155, WAC09436, WAC08854, WAC08921), and Fidaxomicin (Tiacumicin B, WAC02994) summarized in **Supplementary Table 1**.

### **Chaxamycin D purification**

WAC09155 spores were used to inoculate a 50mL of SAM culture which was grown for five days at 30°C and 250rpm. Fermentations were grown in 24 well blocks (EnzyScreen), each containing 3 mL of SM1 media (10g Arkasoy soya flour, 18g glucose, 1g Na<sub>2</sub>SO<sub>4</sub>, and 0.2g CaCO<sub>3</sub> per liter) and 0.3mL of starter culture. Attempts to move to larger fermentation volumes led to a complete loss in production. In total, we cultured approximately 600mL across multiple blocks. After five days of growth at 30°C and 250rpm, we harvested the supernatant by centrifugation at 8 000 x g for 10 minutes and extracted metabolites by incubating overnight with 5% HP-20 (Diaion) resin at 4°C with

stirring. The resin was collected and washed extensively (>10 column volumes) using 5% methanol before elution with 5CV of 95% methanol. The eluted material was dried using rotary evaporation. We removed some impurities by extracting this solid material with hexane/ethyl acetate (3:1) before solubilization in ethyl acetate/methanol (1:1). This material was dried down and resuspended in 2mL of methanol/chloroform (70:30) and separated using a LH-20 (Sephadex) column (3 x 40 cm) equilibrated in methanol. The flow rate was approximately 0.5 mL/min, and we began collecting after ~80 mL, in 3mL fractions. Active fractions were identified by induction of *S. venezuelae*  $\Delta helR\Delta rox$  pGUS: $P_{rox}$ , pooled and dried before further fractionation using reverse phase flash chromatography. The dry material was resuspended in 3mL of methanol and loaded onto a 100g RediSep®Rf Gold C18 column (Teledyne) and fractionated using a CombiFlash (Teledyne) with the following gradient applied at 35mL/min: 0-5 min 10% B; 5-23 min 10-100% B; 23-26.5 min 100%B; 26.5-30min 50% B. Mobile phase A being water + 0.1% formic acid, and B was acetonitrile. Active compounds eluted between 65-85%B, pooled and concentrated before the final purification using reverse phase HPLC (Agilent 1260). We used an X-Select CSH™ prep C18 column (10 x 100 mm, 5  $\mu$ m; Waters) running the following method: 0-1 min 38%B, 1-25 min linear gradient from 38-95%B, 25-27 min 95%B, 27-28 min linear gradient 95-38%B, 28-30 min 38%B. Solvent A was water + 0.1% formic acid Solvent B was acetonitrile. Chaxamycin D eluted at ~62%B at 11min. Chaxamycin D was isolated as a fluffy pale-yellow powder after lyophilization, yielding 0.9mg.

### **Compound 727 purification**

WAC02994 was grown on Bennetts plates for five to seven days at 30 °C until 2-3mm colonies had formed. Two to three colonies were picked into two 3mL SAM cultures containing two 3mm glass beads and incubated for five days at 30°C with shaking (250rpm). The entirety of two of these cultures was used to inoculate 600mL of SAM prepared in a 3L Erlenmeyer flask containing a steel spring for aeration. This starter was grown for five days at 30°C and 250rpm. Fermentations were initiated by adding a 1:20 inoculum of starter culture to 12x 600mL (7.2L total) of SM23 media (11.5g L-proline, 0.5g NaCl, 0.5g K<sub>3</sub>HPO<sub>4</sub>, 30g MOPS with 10mL of 0.2M MgSO<sub>4</sub>, 10mL 0.2M CaCl<sub>2</sub>.and 5mL of Trace salts No 1 added per liter (Trace salts No 1, 1mL 1M H<sub>2</sub>SO<sub>4</sub>, 860mg ZnSO<sub>4</sub>\*7H<sub>2</sub>O, 223mg MnSO<sub>4</sub>\*4H<sub>2</sub>O, 62mg H<sub>3</sub>BO<sub>3</sub>, 125mg CuSO<sub>4</sub>\*5H<sub>2</sub>O, 48mg Na<sub>2</sub>MoO<sub>4</sub>\*2H<sub>2</sub>O, 48mg CoCl<sub>2</sub>\*6H<sub>2</sub>O, 1.8g FeSO<sub>4</sub>\*7H<sub>2</sub>O, and 83mg KI per liter)). SM23 produced the best activity against *E. coli* ΔΔ pGDP3:*rph* out of the media panel described in Zettler *et al.* 2014<sup>62</sup>, a vast improvement over Bennetts and ISP2 (Yeast extract 4g, Malt Extract 10g, Dextrose 4g, per liter, pH 6.8). After five days, we removed the cells by centrifugation (8 000 x g, 15 min) and extracted the supernatant with two volumes of ethyl acetate which were dried by rotary evaporation to afford a crude extract. Compound 727 was isolated from this crude extract in three steps: a normal phase followed by two reverse phase separations. For the first step, the dried extract was reconstituted with 20 mL of chloroform and centrifuged for 10 min at 10 000 x g to pellet insoluble compounds. The supernatant was decanted and separated using a Teledyne CombiFlash® low-pressure chromatography system. A 24 g RediSep® Rf normal-phase silica column was used as a

stationary phase, and the mobile phase consisted of chloroform (A) and methanol (B). A 23-min separation with a flow rate of 35 mL/min was used for the following method: the column was equilibrated to 100% A, the sample was injected, and A was held at 100% for 5 min. Solvent B was increased to 20% over 13 min and held there for 3 more min, then brought back to 0% and held for 2 min. Fractions were collected by 10 mL volume throughout the entire method unless a signal was detected using  $A_{254\text{nm}}$  with a threshold of 0.2 absorbance units. Fractions that had activity against *E. coli*  $\Delta\Delta$  pGDP3(-) and *E. coli*  $\Delta\Delta$  pGDP3:*rph* but not *E. coli*  $\Delta\Delta$  pGDP3:*arr* were pooled together and dried. The pooled fractions were reconstituted in 1 mL of 50% methanol:water v:v and purified by reverse phase HPLC, performed on an Agilent 1260 system. An X-Select CSH™ prep C18 column (10 x 100 mm, 5  $\mu\text{m}$ ; Waters) maintained at 40 °C was used, with a mobile phase of HPLC grade water + 0.1% FA (v/v) (A) and acetonitrile + 0.1% FA (v/v) (B). The injection volume and flow rate were 100  $\mu\text{L}$  and 4 mL/min, respectively. The gradient was held at 10% B for 3 min, increased to 85% B over 13 min, increased to 100% B over 0.5 min, held at 100% B for 3.5 min, then returned to 10% B over 0.5 min. The mobile phase is then held at 10% B for 1.5 min with a 5 min post run time holding at 10% B. The analytes were monitored and collected based on absorbance signals at 254 nm. The active compound had a retention time of 13.2 min and was screened on an Agilent 6546 qToF coupled to a 1290 HPLC LC-MS system. The active compound was determined to have a  $m/z$  of 726.2791 [M-H]<sup>-</sup>. This co-eluted from the first round of HPLC with an inactive compound of  $m/z$  708.2606. These compounds were separated from each other by an additional HPLC fractionation. The stationary and mobile phases were the same as listed above but with a

different gradient. The final round of isolation had a 40% B starting and finishing concentrations with identical time and concentration ramp steps as above. The RT of compound 727 was 14.87 min with a final yield of <0.5 mg.

### **Antimicrobial susceptibility testing**

The standard CLSI broth microdilution method was used to determine minimum inhibitory concentrations in cation-adjusted Mueller Hinton broth (CAMHB, BD) and incubation at 37°C. All reported MICs were determined from two independent experiments conducted in technical duplicates. Bioassays used during rifamycin dereplication (**Figure 3**), activity-guided purification (**Figure 4C**), and the synergy assay (**Figure 4A**) were performed in the same manner but used LB in place of CAMHB.

### **Genome sequencing**

WAC strains were grown for 24-72 h in TSB medium at 30°C, 250rpm, and mycelium was harvested from mid-late log phase cultures. Genomic DNA was isolated using the Wizard<sup>®</sup> Genomic DNA Purification Kit (Promega) according to manufacturer instructions for Gram positive bacteria, and quality was assessed by spectrophotometry and gel electrophoresis. Sequencing libraries were prepared for paired-end reads using the NEB Next Ultra V1 kit with 500ng of DNA input (sonicated to 800 bp), followed by two size selections using AMPure XP beads (Beckman Coulter). Sonication, size selection, and ultimately Illumina sequencing (MiSeq 2 x 250bp reads) was performed by the McMaster Genomics Facility in the Farncombe Institute (Hamilton, ON, Canada). Raw reads were trimmed and merged using skewer v0.2.2 and FLASH v1.2.11, respectively, and *de novo*

assemblies were built using SPAdes v3.11.1<sup>63-65</sup>. Assemblies were then analyzed for biosynthetic gene clusters using antiSMASH v6.0<sup>66</sup>.

### **Analysis of rifamycin-like biosynthetic gene clusters**

We used cblaster to query all Actinobacterial genomes in RefSeq using a minimal suite of genes common to all naphthlenic rifamycins<sup>67</sup>. These genes were *rifGHJKLM* (The full set of AHBA synthesis genes), *rifF* (Amide synthase/cyclase), and *orf19* (hydroxylase responsible for the formation of the second 6-membered ring fused with AHBA), all taken from the *Amycolatopsis mediterranei* S699 cluster (Genbank acc. AF040570.3)<sup>52</sup>. Within our cblaster query we used the following specific cutoffs, hits require a minimum 55% identity at the amino acid level, hits can be no farther than 80kb apart (necessary to identify genes on either side of the PKS genes), candidate clusters must contain *rifG*, *rifK*, and *orf19*. This identified 125 clusters; we downloaded the flanking 100kb up and downstream regions for each hit and extracted the BGC by analyzing each sequence using antiSMASH v6.0<sup>66</sup>. From antiSMASH, we obtained 84 complete and partial ansamycin clusters; we extracted the sequence of each *rifK* homolog from this dataset, as well as the seven rifamycin-like clusters obtained in this work, and aligned them using Muscle v5.1 and constructed a phylogenetic tree using the WAG substitution model on FastTree v2.1.11 with default settings, iTOLv5.0 was used to visualize the tree and prepare **Figure 5**, **Supplementary Figure 7**<sup>68-70</sup>. Clinker was used to compare clusters, color-code genes by function, and create cluster diagrams featured in **Figure 5 and Supplementary Figure 3**<sup>71</sup>. RpoB sequences were extracted from each naphthlenic ansamycin encoding genome using the TIGRFAM model of bacterial DNA-dependent RNA polymerase subunit beta

(TIGR02013.1) as a query for HMMER3<sup>72</sup>. Sequences were aligned using Muscle v5.1 and manually extracted the four regions associated with rifamycin resistance mutations, rif clusters I-III and N (**Supplementary Figure 7**)<sup>30,73</sup>.

### **WAC09654 cluster capture**

We used transformation-assisted recombination (TAR) in *S. cerevisiae* VL6-48N to capture the rifamycin-like cluster from WAC09654. This procedure has been detailed in several previous publications from our group, so we will not fully reproduce it here<sup>60,74–76</sup>. A gblock containing our 50bp capture hooks separated by a *PmeI* site for linearization of the capture vector was cloned into *NdeI/XhoI* digested pCGW using Gibson assembly. The sequence for the gblock is as follows (homology arms are underlined, *PmeI* site in bold).

GCCTCCCATGGTATAAATAGTGGCAGTCCGGTGAAGTTCGTCCCGTTCGAGG  
GCAAGGCGATGGTCATGGGCCTGTTTAAACGAGGGGTCCGACGCGGTGGTC  
ATGGAGCCAGCCTAAGGGCAGGTTCCGGCATGTTCGAAAGCTACATATAAGG

Sequence verified pCGW containing our capture arms was linearized with *PmeI* and co-transformed into *S. cerevisiae* spheroplasts alongside gDNA from WAC09654 digested with *Psp1406I* and *XbaI*, which cut near the ends of the cluster but not within it (**Supplementary Figure 5B**). HMW gDNA was obtained through in-gel lysis and digestion of agarose plugs containing WAC09654, as described in Kieser *et al.* 1992<sup>77</sup>. Yeast transformants could be observed after ~5 days of incubation at 30°C, gDNA was extracted from transformants and screened by PCR for three different regions of the cluster (**Supplementary Figure 5A and C**) using the following primers cPCR1 FP 5'-GATGACGTCCGCCTCTACAG-3', RP 5'-AGAGCGGAGTACCAGACCAT-3', cPCR2 FP 5'-CCGATCCAACGTAGGTCTCG-3', RP 5'-

GACGAGATCGCCGGATTCTT-3', cPCR3 FP 5'-AAGTAGTTGGTGCCTCGAA-3', RP 5'-GTGTCCTCGTCGATCGACTC-3'. Clones testing positive for all three were transformed into *E. coli* Epi300 to recover the plasmid and produce enough for restriction mapping to confirm our cluster's capture (**Supplementary Figure 5D**). We then transformed pCGW:9654 into *E. coli* ET12457 pUZ8002::*bla* and used biparental mating to introduce it into *S. coelicolor* M1154 as per well-established protocols<sup>78</sup>.

The LuxR family regulator was amplified from WAC09654 gDNA using the following primers FP 5'-AACCCATGCATATGCCCGATTCCGTGACTCGC-3' and RP 5'- CCACACTCGAGTCAGCCGGCGTTCGCGAG3' and cloned into pIJ10257 downstream of the constitutive  $P_{ermE^*}$  promoter using *NdeI* and *XhoI* (sequences underlined). After confirming the correct insert sequence, we conjugated pIJ:*luxR* into M1154 and M1154 pCGW:9654 using *E. coli* ET12457 pUZ8002::*bla* as described in Kieser *et al.* 2001<sup>78</sup>.

### **Heterologous expression of WAC09654**

Spores of *S. coelicolor* M1154 were inoculated into 3mL of SAM media with one sterile glass bead and incubated for 48 hours at 30°C and 250rpm. 50mL of Bennett's media or Bennett's media + 80mM KNO<sub>3</sub> were inoculated with a 1:20 dilution of SAM starter culture and incubated for five days at 30°C and 250rpm. All fermentations were carried out in 250mL flasks with steel spring baffles for aeration and homogenization. After five days of growth, cells were removed by centrifugation at 4 000 x g for 15 min, and the supernatant was extracted twice with ethyl acetate and dried using a centrifugal evaporator

(Genevac/SP Scientific). The dried material was resuspended in 500  $\mu$ L of DMSO, used to assess RAE activation, and analyzed by LC-MS.

## REFERENCES

1. World Health Organization. *Global antimicrobial resistance and use surveillance system (GLASS) report 2021*. World Health Organization. (2021).
2. O'Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. London: Review on Antimicrobial Resistance, 2016. doi:10.4103/2045-080x.186181
3. *2020 Antibacterial Agents in Clinical and Preclinical Development: an overview and analysis*. Geneva: World Health Organization; 2021.
4. Centers for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States, 2019*. Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2019.
5. Laxminarayan, R. *et al.* Antibiotic resistance-the need for global solutions. *Lancet Infect. Dis.* **13**, 1057–1098 (2013).
6. Lewis, K. The Science of Antibiotic Discovery. *Cell* **181**, 29–45 (2020).
7. Theuretzbacher, U., Outterson, K., Engel, A. & Karlén, A. The global preclinical antibacterial pipeline. *Nat. Rev. Microbiol.* **18**, 275–285 (2020).
8. Butler, M. S. *et al.* Analysis of the Clinical Pipeline of Treatments for Drug-Resistant Bacterial Infections: Despite Progress, More Action Is Needed. *Antimicrob. Agents Chemother.* **66**, 1–20 (2022).
9. Cox, G. *et al.* A Common Platform for Antibiotic Dereplication and Adjuvant Discovery. *Cell Chem. Biol.* **24**, 98–109 (2017).
10. Baltz, R. H. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J. Ind. Microbiol. Biotechnol.* **33**, 507–513 (2006).
11. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
12. Gavriilidou, A. *et al.* Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes. *Nat. Microbiol.* **7**, 726–735 (2022).
13. Xu, M. & Wright, G. D. Heterologous expression-facilitated natural products' discovery in actinomycetes. *J. Ind. Microbiol. Biotechnol.* **46**, 415–431 (2019).
14. Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523 (2015).
15. Lee, N. *et al.* Systems and synthetic biology to elucidate secondary metabolite biosynthetic gene clusters encoded in: *Streptomyces* genomes. *Nat. Prod. Rep.* **38**, 1330–

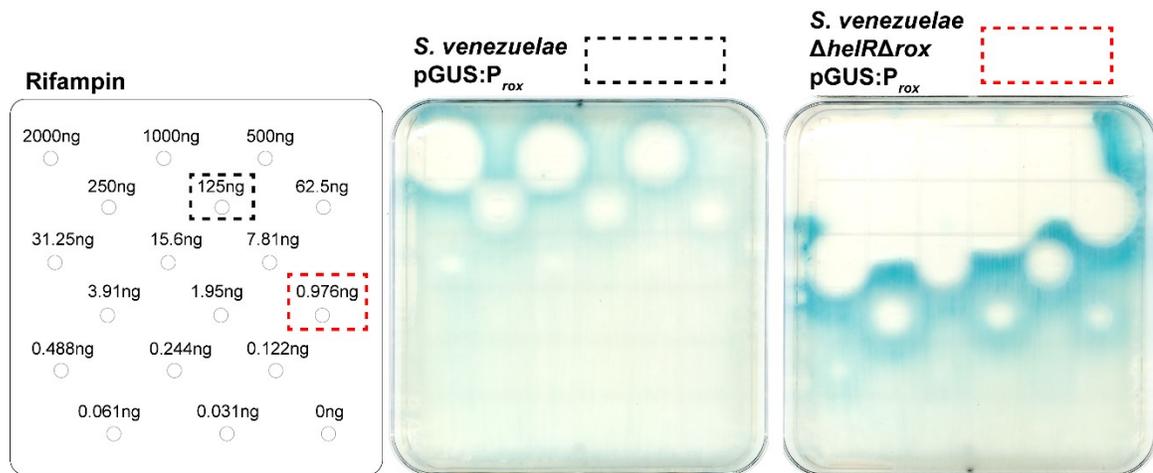
- 1361 (2021).
16. Fischer, H. P. *et al.* Identification of antibiotic stress-inducible promoters: A systematic approach to novel pathway-specific reporter assays for antibacterial drug discovery. *Genome Res.* **14**, 90–98 (2004).
  17. Kurita, K. L., Glassey, E. & Linington, R. G. Integration of high-content screening and untargeted metabolomics for comprehensive functional annotation of natural product libraries. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11999–12004 (2015).
  18. Montano, E. T. *et al.* Bacterial Cytological Profiling Identifies Rhodanine-Containing PAINS Analogs as Specific Inhibitors of *Escherichia coli* Thymidylate Kinase In Vivo. *J. Bacteriol.* **203**, 105–121 (2021).
  19. Wex, K. W. *et al.* Bioreporters for direct mode of action-informed screening of antibiotic producer strains. *Cell Chem. Biol.* **28**, 1242-1252.e4 (2021).
  20. Gallagher, L. A., Bailey, J. & Manoil, C. Ranking essential bacterial processes by speed of mutant death. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 18010–18017 (2020).
  21. Silvis, M. R. *et al.* Morphological and transcriptional responses to CRISPRi knockdown of essential genes in *Escherichia coli*. *MBio* **12**, (2021).
  22. Ma, C., Yang, X. & Lewis, P. Bacterial Transcription as a Target for Antibacterial Drug Development. *Microbiol. Mol. Biol. Rev.* **80**, 139–160 (2016).
  23. Kirsch, S. H., Haeckl, F. P. J. & Müller, R. Beyond the approved : target sites and inhibitors of bacterial RNA polymerase from bacteria and fungi. *Nat. Prod. Rep.* **39**, 1226–1263 (2022).
  24. Boyaci, H. *et al.* Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpAcontacts. *Elife* **7**, 1–19 (2018).
  25. Lin, W. *et al.* Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Mol. Cell* **70**, 60–71.e15 (2018).
  26. Campbell, E. A. *et al.* Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **104**, 901–912 (2001).
  27. Feklistov, A. *et al.* Rifamycins do not function by allosteric modulation of binding of Mg<sup>2+</sup> to the RNA polymerase active center. *Proc. Natl. Acad. Sci.* **105**, 14820–14825 (2008).
  28. Rothstein, D. M. Rifamycins, alone and in combination. *Cold Spring Harb. Perspect. Med.* **6**, 1–19 (2016).
  29. Goldstein, B. P. Resistance to rifampicin: a review. *J. Antibiot. (Tokyo)*. **67**, 625–630 (2014).
  30. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber. Lung Dis.* **79**, 3–29 (1998).
  31. Peek, J. *et al.* Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nat. Commun.* **9**, (2018).

32. Mosaei, H. *et al.* Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant *Mycobacterium tuberculosis*. *Mol. Cell* **72**, 263–274.e5 (2018).
33. Surette, M. D., Spanogiannopoulos, P. & Wright, G. D. The Enzymes of the Rifamycin Antibiotic Resistome. *Acc. Chem. Res.* **111**, (2021).
34. Surette, M. D., Waglechner, N., Koteva, K. & Wright, G. D. HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics. *Mol. Cell* **82**, 3151–3165.e9 (2022).
35. Rominski, A., Roditscheff, A., Selchow, P., Böttger, E. C. & Sander, P. Intrinsic rifamycin resistance of *Mycobacterium abscessus* is mediated by ADP-ribosyltransferase MAB\_0591. *J. Antimicrob. Chemother.* dkw466 (2016). doi:10.1093/jac/dkw466
36. Spanogiannopoulos, P., Waglechner, N., Koteva, K. & Wright, G. D. A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7102–7107 (2014).
37. Ma, Z. & Lynch, A. S. Development of a Dual-Acting Antibacterial Agent (TNP-2092) for the Treatment of Persistent Bacterial Infections. *J. Med. Chem.* **59**, 6645–6657 (2016).
38. Betina, V., Micekova, D. & Nemeč, P. Antimicrobial Properties of Cytochalasins and Their Alteration of Fungal Morphology. *J. Gen. Microbiol.* **71**, 343–349 (1972).
39. Bottalico, A., Capasso, R., Evidente, A., Randazzo, G. & Vurro, M. Cytochalasins: Structure-activity relationships. *Phytochemistry* **29**, 93–96 (1990).
40. Rateb, M. E. *et al.* Chaxamycins A - D, bioactive ansamycins from a hyper-arid desert *Streptomyces* sp. *J. Nat. Prod.* **74**, 1491–1499 (2011).
41. Xiao, Y. S. *et al.* Rifamorpholines A–E, potential antibiotics from locust-associated actinobacteria *Amycolatopsis* sp. Hca4. *Org. Biomol. Chem.* 3909–3916 (2017). doi:10.1039/c7ob00614d
42. Williams, D. E. *et al.* Aminorifamycins and Sporalactams Produced in Culture by a *Micromonospora* sp. Isolated from a Northeastern-Pacific Marine Sediment Are Potent Antibiotics. *Org. Lett.* **19**, 766–769 (2017).
43. Thaker, M. N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
44. Wu, Y., Kang, Q., Shen, Y., Su, W. & Bai, L. Cloning and functional analysis of the naphthomycin biosynthetic gene cluster in *Streptomyces* sp. CS. *Mol. Biosyst.* **7**, 2459–2469 (2011).
45. Liu, Y. *et al.* Functional Analysis of Cytochrome P450s Involved in Streptovaricin Biosynthesis and Generation of Anti-MRSA Analogues. *ACS Chem. Biol.* **12**, 2589–2597 (2017).
46. Castro, J. F. *et al.* Identification and heterologous expression of the chaxamycin biosynthesis gene cluster from *Streptomyces leeuwenhoekii*. *Appl. Environ. Microbiol.* **81**, 5820–5831 (2015).

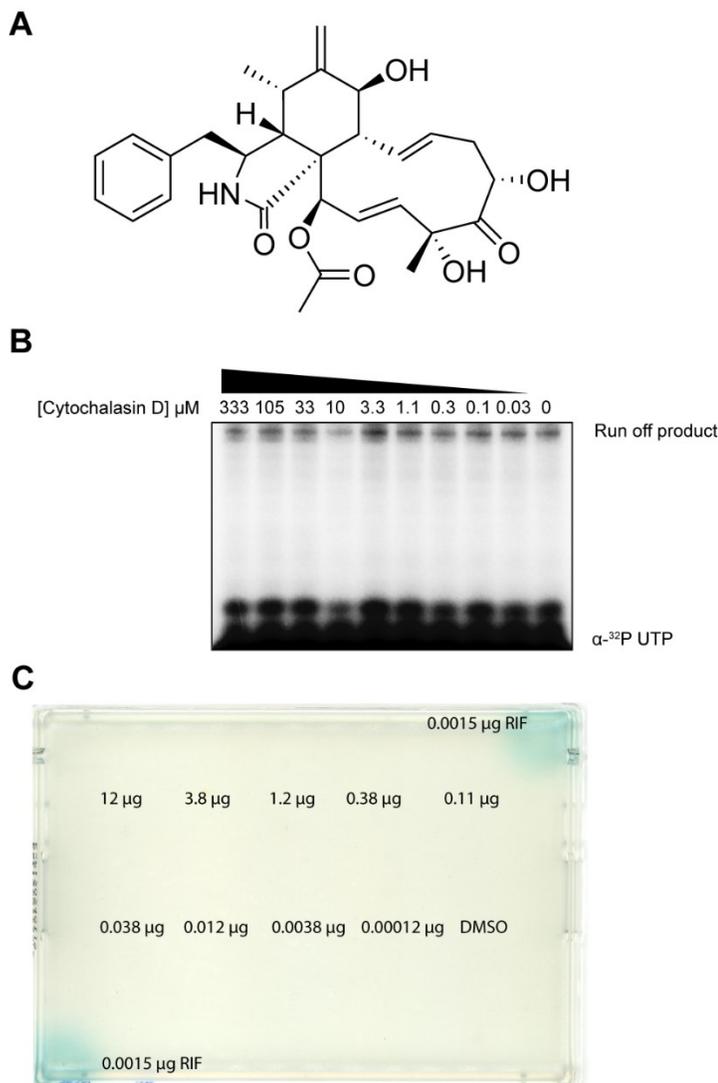
47. Barker, R. D., Yu, Y., De Maria, L., Johannissen, L. O. & Scrutton, N. S. Mechanism of Action of Flavin-Dependent Halogenases. *ACS Catal.* **12**, 15352–15360 (2022).
48. Sun, G. *et al.* Uncovering the cytochrome P450-catalyzed methylenedioxy bridge formation in streptovaricins biosynthesis. *Nat. Commun.* **11**, 1–12 (2020).
49. Gomez-Escribano, J. P. & Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol.* **4**, 207–215 (2011).
50. Li, S. *et al.* Activating a Cryptic Ansamycin Biosynthetic Gene Cluster to Produce Three New Naphthalenic Octaketide Ansamycins with n-Pentyl and n-Butyl Side Chains. *Org. Lett.* **17**, 3706–3709 (2015).
51. Shao, H. Z. *et al.* A preliminary study of the mechanism of nitrate-stimulated remarkable increase of rifamycin production in *Amycolatopsis mediterranei* U32 by RNA-seq. *Microb. Cell Fact.* **14**, 1–11 (2015).
52. Heinz G. Floss & Yu, T.-W. Rifamycin-Mode of Action, Resistance, and Biosynthesis. *Chem. Rev.* **105**, 621–632 (2005).
53. Kim, H. *et al.* Mechanism of natural rifampin resistance of *Streptomyces* spp. *Syst. Appl. Microbiol.* **28**, 398–404 (2005).
54. Newell, K. V, Thomas, D. P., Brekasis, D. & Paget, M. S. B. The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on *Streptomyces coelicolor*. *Mol. Microbiol.* **60**, 687–696 (2006).
55. MacNeil, D. J. *et al.* Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68 (1992).
56. Koteva, K. *et al.* Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem. Biol.* **25**, 403–412.e5 (2018).
57. Noskov, V. N. *et al.* A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* **4**, 1–11 (2003).
58. Spanogiannopoulos, P., Thaker, M., Koteva, K., Waglechner, N. & Wright, G. D. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob. Agents Chemother.* **56**, 5061–5069 (2012).
59. Hong, H. J., Hutchings, M. I., Hill, L. M. & Buttner, M. J. The role of the novel fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J. Biol. Chem.* **280**, 13055–13061 (2005).
60. Xu, M. *et al.* GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. *Nat. Commun.* **11**, (2020).
61. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
62. Zettler, J. *et al.* New Aminocoumarins from the rare actinomycete *Catenulispora acidiphila* DSM 44928: Identification, structure elucidation, and heterologous production.

- ChemBioChem* **15**, 612–621 (2014).
63. Nurk, S. *et al.* *Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads In: Research in Computational Molecular Biology. RECOMB 2013. Lecture Notes in Computer Science* (2013).
  64. Magoč, T. & Salzberg, S. L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
  65. Jiang, H., Lei, R., Ding, S. W. & Zhu, S. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 1–12 (2014).
  66. Blin, K. *et al.* AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **49**, W29–W35 (2021).
  67. Gilchrist, C. L. M. *et al.* Cblaster: a Remote Search Tool for Rapid Identification and Visualization of Homologous Gene Clusters. *Bioinforma. Adv.* **1**, 1–10 (2021).
  68. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
  69. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, (2010).
  70. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 0–1 (2022).
  71. Gilchrist, C. L. M. & Chooi, Y. H. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473–2475 (2021).
  72. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, (2011).
  73. Jin, D. J. & Gross, C. A. Mapping and sequencing of mutations in the *Escherichia coli* rpoB gene that lead to rifampicin resistance. *J. Mol. Biol.* **202**, 45–58 (1988).
  74. Tang, X. *et al.* Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-direct Genome Mining. *ACS Chem. Biol.* **10**, 2841–2849 (2015).
  75. Xu, M. *et al.* Phylogeny-Informed Synthetic Biology Reveals Unprecedented Structural Novelty in Type v Glycopeptide Antibiotics. *ACS Cent. Sci.* **8**, 615–626 (2022).
  76. Culp, E. J. *et al.* ClpP inhibitors are produced by a widespread family of bacterial gene clusters. *Nat. Microbiol.* **7**, 451–462 (2022).
  77. Kieser, H. M., Kieser, T. & Hopwood, D. A. A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. *J. Bacteriol.* **174**, 5496–5507 (1992).
  78. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. & Hopwood, D. *Practical Streptomyces Genetics*. (John Innes Foundation, 2000).
  79. Terlouw, B. R. *et al.* MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res.* **51**, D603–D610 (2023).

### Supplementary Materials



**Supplementary Figure 1 Hypersusceptible *S. venezuelae* pGUS:P<sub>rox</sub> improves detection of rifampin significantly.** Comparison of RAE induction (blue) by lawns of wildtype *S. venezuelae* pGUS:P<sub>rox</sub> and a double knockout of the RNA polymerase protection protein HelR and the rifamycin monooxygenase Rox exposed to various quantities of rifampin. *S. venezuelae* Δ*helR*Δ*rox* is ~64 fold more susceptible to rifampin, and the limit of detection is ~128 fold lower.



**Supplementary Figure 2 Cytochalasin D does not inhibit RNAP or induce the RAE.**

**A) Structure of Cytochalasin D.** **B) *In vitro* transcription reactions containing cytochalasin D.** Reactions were performed using *S. venezuelae* RNAP and  $\sigma^{\text{HrdB}}$  with  $P_{AP3}$  as template as described previously<sup>34</sup>. Reaction products were separated on a 20% PAGE gel containing 7M urea and visualized by autoradiography. Reactions contained  $\alpha^{32}\text{P}$  UTP to enable visualization of the full length transcript (run-off product). **C) Cytochalasin D does not activate induction of the RAE.** SIM agar plate inoculated with *S. venezuelae*  $\Delta\text{helR}\Delta\text{rox}$  pGUS: $P_{\text{rox}}$  imaged 72 hours after cytochalasin D was applied to the plate in the quantities indicated. 1.5 ng of rifampin was added to the corners of the plate as a positive control for induction. All compounds were dissolved in DMSO.

**Supplementary Table 1 – Additional hit information**

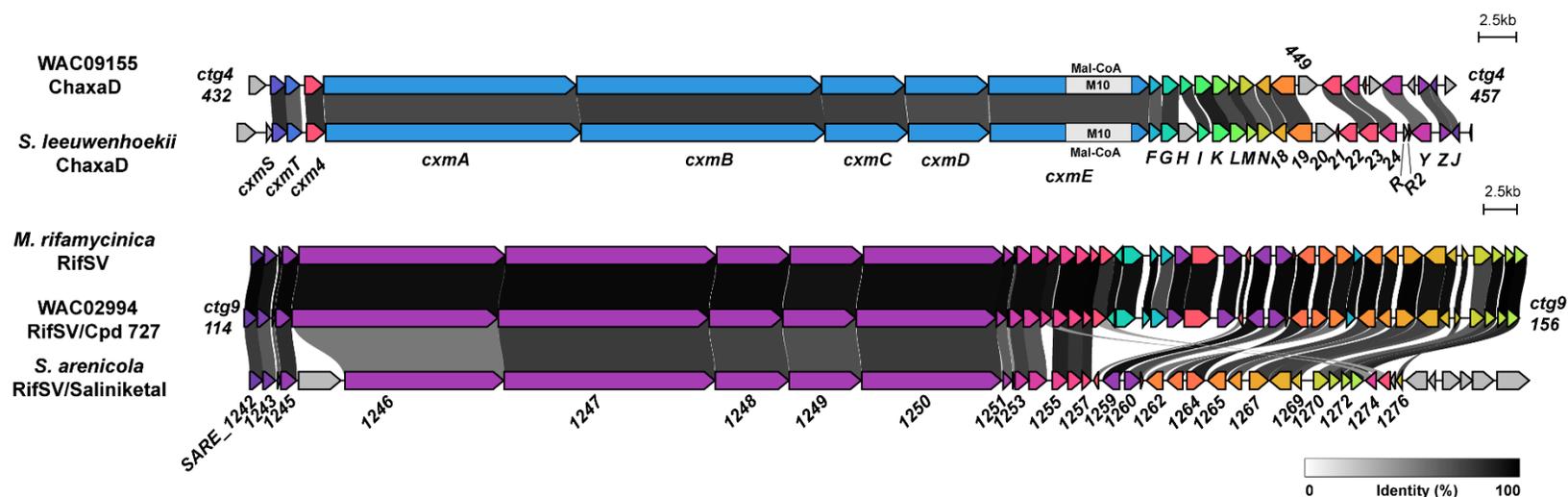
<b>Strain</b>	<b>Replicated</b>	<b>Clonal<sup>1</sup></b>	<b>LC-MS*/GNPS<sup>2</sup></b>	<b>Antismash<sup>3</sup></b>
<b>WAC09654</b>	<b>Yes</b>		<b>None</b>	<b>Rif-like cluster</b>
<b>WAC00728</b>	<b>Yes</b>		None	N/A
<b>WAC08122</b>	<b>Yes</b>		None	N/A
<b>WAC02994</b>	<b>Yes</b>	<b>Yes</b>	<b>Rifamycin SV/ Fidaxomicin</b>	<b>Rifamycin SV/ Fidaxomicin</b>
WAC02993		02994		
<b>WAC10744</b>	<b>Yes</b>		<b>Rifamycin SV/ Rifamycin B*</b>	<b>Rifamycin SV</b>
<b>WAC09155</b>	<b>Yes</b>		<b>Chloramphenicol/ Chaxamycin D</b>	<b>Chaxamycin/ Chloramphenicol</b>
WAC09157		09155		
WAC09165	Yes	09155		
WAC09167		09155		
WAC08854	Yes	09155?	Chloramphenicol/ Chaxamycin D*	
WAC09436	Yes	09155	Chloramphenicol/ Chaxamycin D*	
WAC08921	Yes	09155	Chloramphenicol/ Chaxamycin D*	
WAC08922		09155		
<b>WAC06666</b>	<b>Yes</b>		<b>Rifamycin SV</b>	<b>Rifamycin SV</b>
WAC06667		06666		
WAC06669		06666		
<b>WAC07128</b>	<b>Yes</b>		<b>Rifamycin SV</b>	<b>Rifamycin SV</b>
<b>WAC03369</b>	<b>Yes</b>		<b>None</b>	<b>Rif-like cluster</b>
WAC06439	No			
WAC04336	No			
WAC06344	No			
WAC08454	No			
WAC09446	No			
WAC06248	No			
WAC08802	N/A			<b>Rifamycin SV</b>
<b>Known producers</b>				
WAC07295			Rifamycin SV*	Rifamycin SV
WAC06771			Rifamycin SV*	Rifamycin SV
WAC07376			Rifamycin SV*	Rifamycin SV
WAC06369			Rifamycin SV*	Rifamycin SV

\*Identified based on HRMS, not GNPS specifically

<sup>1</sup>Clonal isolates were from the same soil sample, and appeared identical morphologically.

<sup>2</sup>GNPS hits to library standards were based on a default settings<sup>61</sup>.

<sup>3</sup>Cluster predictions by AntiSMASHv6.0, only BGCs relevant to the bioactivity in **Figure 5** are noted here



**Supplementary Figure 3 BGC similarity of WAC09155 and WAC02994 to BGCs with known products.** Conserved proteins between clusters are signaled by connecting bars, which are darker for higher % identity at the amino acid level. The WAC09155 BGC is highly similar to the documented chaxamycin D (ChaxaD) BGC from *S. leeuwenhoekii* (LN831790). Importantly, it is predicted to accept a malonyl-CoA (instead of methylmalonyl-CoA) in its 10<sup>th</sup> module, just like chaxamycins. The WAC02994 BGC is nearly identical to a cluster from *Micromonospora rifamycinica* (LT607752), known to produce rifamycin SV. A second comparison was performed to the Rifamycin SV/Saliniketal cluster from *Salinispora arenicola* (NZ\_KB913036) as this cluster has been studied in more detail. Protein names are given for the well-characterized clusters. Genes are colored by similarity, but only within the clusters they were directly compared to (WAC09155 v *S. leeuwenhoekii* colors do not correspond to WAC02994 v *M. rifamycinica* v *S. arenicola*). This figure was generated using clinker.

Supplementary Table 2 – WAC09155 BGC comparison

Chaxamycin BGC*	WAC90155 BGC	Identity (%)**	Similarity (%)
CxmZ	ctg4_457	0.73	0.78
CxmZ	ctg4_456	0.76	0.82
CxmY	ctg4_454	0.66	0.77
CxmR2			
CxmR			
Cxm24	ctg4_451	0.73	0.8
Cxm23	ctg4_450	0.75	0.83
Cxm22			
Cxm21			
	ctg4_449***		
Cxm20***	N/A		
Cxm19	ctg4_448	0.8	0.86
Cxm18	ctg4_447	0.77	0.83
CxmN	ctg4_446	0.72	0.83
CxmM	ctg4_445	0.85	0.91
CxmL	ctg4_444	0.78	0.86
CxmK	ctg4_443	0.9	0.94
CxmI	ctg4_442	0.82	0.88
CxmG	ctg4_441	0.82	0.85
CxmF	ctg4_440	0.72	0.82
CxmE	ctg4_439	0.77	0.84
CxmD	ctg4_438	0.76	0.85
CxmC	ctg4_437	0.8	0.87
CxmB	ctg4_436	0.81	0.88
CxmA	ctg4_435	0.78	0.85
Cxm4	ctg4_434	0.84	0.9
CxmT	ctg4_433	0.71	0.81
CxmS	ctg4_432	0.85	0.91

\**S. leeuwenhoekii* C34 (Genbank acc. LN831790)

\*\*Calculated at the amino-acid level

\*\*\*Cxm20 and ctg4\_449 both encode acetyltransferases, but 449 is a different family (OafA) from those usually associated with rifamycin BGCs (PapA5)

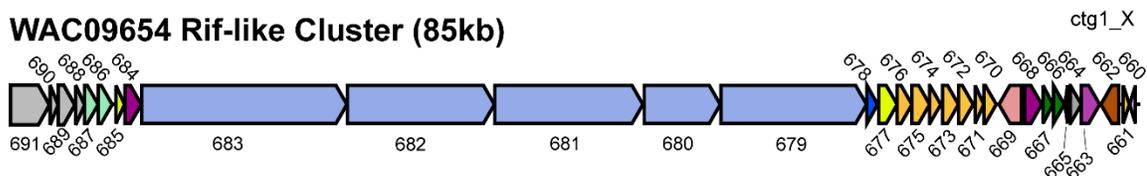
Supplementary Table 3 – WAC02994 BGC comparison

WAC02994	RifSV BGC- <i>Mr</i> *	%ID/SIM	RifSV BGC- <i>Sa</i> **	%ID/SIM
ctg9_156	SCG35264.1	0.98/0.99	SARE_1242	0.81/0.89
ctg9_155	SCG35274.1	0.93/0.95	SARE_1243	0.72/0.82
ctg9_154	SCG35282.1	0.96/0.97	SARE_1244	0.7/0.8
ctg9_153	SCG35291.1	0.98/0.98	SARE_1245	0.86/0.91
ctg9_152	SCG35299.1	0.95/0.97	SARE_1246***	0.61/0.67
ctg9_151	SCG35325.1	0.97/0.98	SARE_1247	0.82/0.88
ctg9_150	SCG35341.1	0.97/0.98	SARE_1248	0.79/0.87
ctg9_149	SCG35350.1	0.94/0.96	SARE_1249	0.75/0.84
ctg9_148	SCG35362.1	0.96/0.97	SARE_1250	0.82/0.88
ctg9_147	SCG35372.1	0.97/0.98	SARE_1251	0.71/0.76
ctg9_146	SCG35380.1	0.76/0.83	SARE_1252	0.5/0.57
ctg9_145	SCG35390.1	0.96/0.98	SARE_1253	0.8/0.87
ctg9_144	SCG35400.1	0.96/0.97	SARE_1254	0.73/0.8
ctg9_143	SCG35410.1	0.94/0.96	SARE_1274	0.52/0.64
ctg9_142	SCG35419.1	0.98/0.98	SARE_1275	0.89/0.94
ctg9_141	SCG35429.1	0.98/0.99	SARE_1255	0.83/0.87
ctg9_140	SCG35440.1	0.99/0.99	SARE_1256	0.87/0.93
ctg9_139	SCG35449.1	0.85/0.89	SARE_1257	0.47/0.59
ctg9_138	SCG35459.1	0.96/0.97		
ctg9_137	SCG35468.1	0.95/0.97		
ctg9_136	SCG35476.1	0.98/0.99		
ctg9_135	SCG35485.1	0.78/0.79		
ctg9_134	SCG35493.1	0.97/0.99		
ctg9_133	SCG35503.1	0.96/0.97		
ctg9_132	SCG35515.1	0.96/0.98		
ctg9_131	SCG35525.1	0.97/0.99	SARE_1258	0.61/0.73
ctg9_129	SCG35537.1	0.33/0.48	SARE_1259	0.94/0.97
ctg9_128	SCG35557.1	0.97/0.98	SARE_1260	0.76/0.83
ctg9_127	SCG35570.1	0.98/0.99	SARE_1261	0.63/0.69
ctg9_126	SCG35582.1	0.97/0.99	SARE_1262	0.89/0.94
ctg9_125	SCG35589.1	0.94/0.97	SARE_1263	0.78/0.86
ctg9_124	SCG35599.1	0.96/0.96	SARE_1264	0.63/0.78
ctg9_123	SCG35608.1	0.98/0.99	SARE_1265	0.8/0.87
ctg9_122	SCG35617.1	0.95/0.97	SARE_1266	0.73/0.79
ctg9_121	SCG35627.1	0.99/0.99	SARE_1267	0.8/0.87
ctg9_120	SCG35638.1	0.94/0.95	SARE_1268	0.79/0.88
ctg9_119	SCG35653.1	0.96/0.97	SARE_1269	0.75/0.84
ctg9_118	SCG35661.1	0.94/0.96	SARE_1270	0.76/0.86
ctg9_117	SCG35671.1	0.78/0.78	SARE_1271	0.7/0.77
ctg9_116	SCG35680.1	0.97/0.98	SARE_1272	0.81/0.89
ctg9_115	SCG35689.1	1	SARE_1273	0.83/0.87
ctg9_114	SCG35698.1	0.98/0.99	SARE_1276	0.88/0.95

\**Micromonospora rifamycinica* (Genbank acc. LT607752)

\*\**Salinispora arenicola* (Genbank acc. NZ\_KB913036)

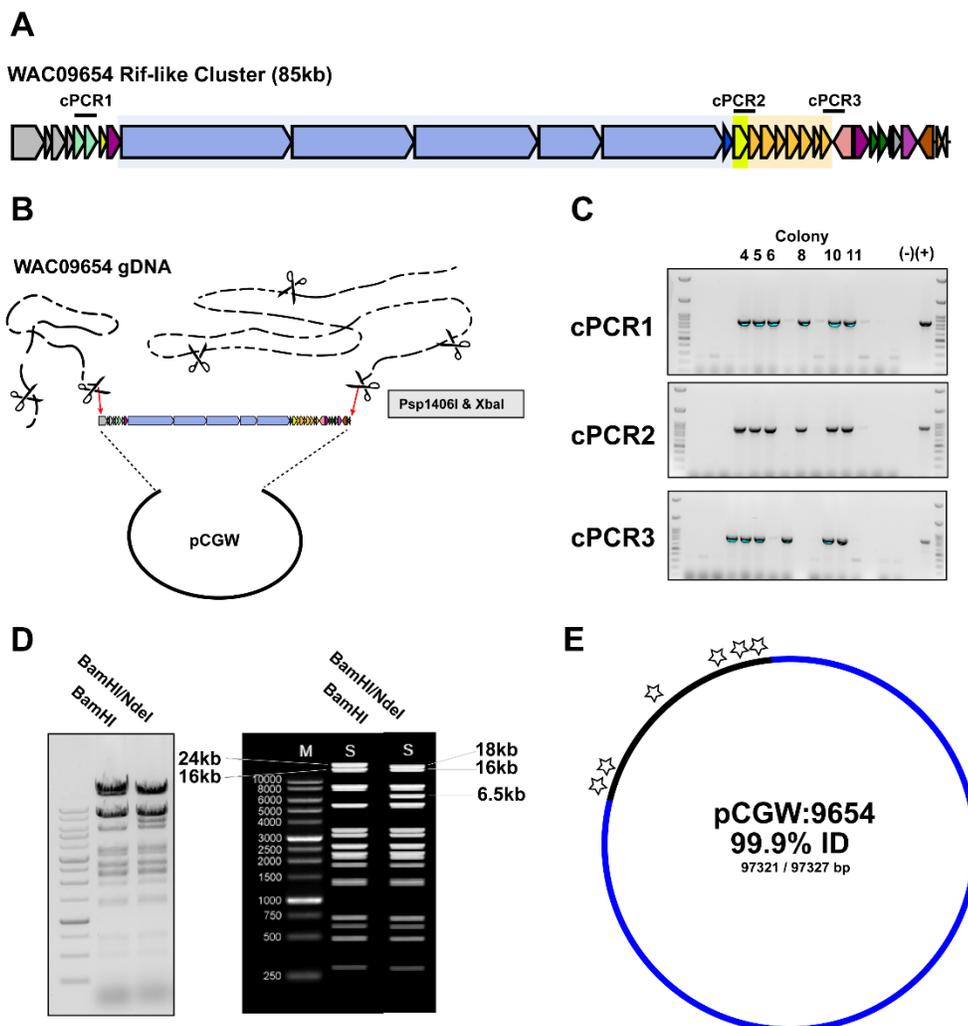
\*\*\*Assembly splits this protein in two, so the identity and similarity are higher than reported here

**WAC09654 Rif-like Cluster (85kb)**

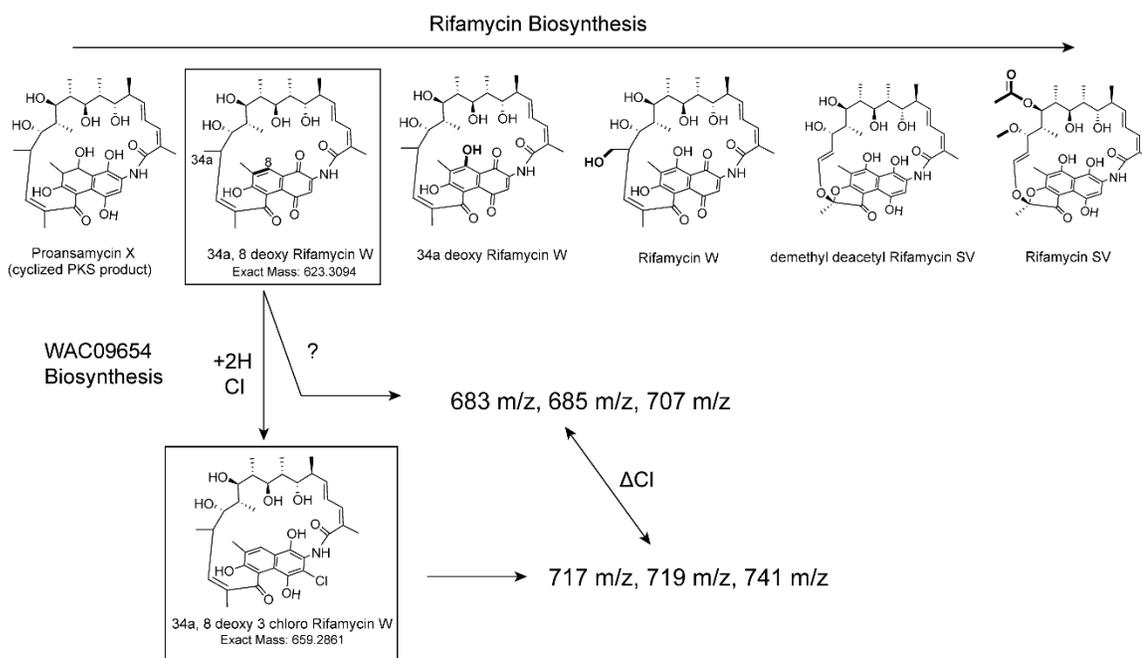
9654rif Locus	Closest MIBIG hit	Function	%ID	Query coverage	Bitscore	E-value
ctg1_660	N/A	Nuclear transport factor 2 family protein	N/A	N/A	N/A	N/A
<b>ctg1_661</b>	ASZ00169.1	StvJ (aminodehydroquinone dehydratase, RifI)	85	99.3	262	5.26E-91
<b>ctg1_662</b>	ASZ00166.1	StvR1 (LuxR family regulator)	70	98.2	563	2.85E-201
<b>ctg1_663</b>	ADM46372.1	Nat4 (membrane ion antiporter)	66	95.5	520	2.22E-183
<b>ctg1_664</b>	ABV97174.1	RifI (Shikimate dehydrogenase)	66	95.1	344	2.57E-120
<b>ctg1_665</b>	ASZ00165.1	ferredoxin	65	96.9	88	4.98E-25
<b>ctg1_666</b>	ASZ00164.1	StvM3 (Methyltransferase)	67	100.4	361	7.38E-127
<b>ctg1_667</b>	CBH32791.1	Methyltransferase/dehydrogenase	52	95.8	216	2.21E-65
<b>ctg1_668</b>	ASZ00162.1	StvP4 (Cytochrome P450)	71	99.3	605	4.11E-218
<b>ctg1_669</b>	ASZ00161.1	StvO4 (hydroxylase, intramolecular ring formation, Orf19)	78	99.8	808	5.60E-294
<b>ctg1_670</b>	CQR60485.1	MitN (kanosamine kinase, RifN)	72	98.7	412	9.38E-146
<b>ctg1_671</b>	AAC01721.1	RifM (Phosphatase)	79	99	326	2.53E-114
<b>ctg1_672</b>	WP_033261451.1	RifL homolog (oxidoreductase)	72	100.3	495	1.12E-176
<b>ctg1_673</b>	ASZ00156.1	StvK (AHBA synthase, RifK)	83	99.7	656	2.20E-239
<b>ctg1_674</b>	ASZ00155.1	StvI (shikimate dehydrogenase, RifI)	80	99.6	425	5.62E-152
<b>ctg1_675</b>	CAI94725.1	RubH (putative aDAMP synthase, RifH)	74	94.2	580	3.79E-208
<b>ctg1_676</b>	ASZ00153.1	StvG (3-dehydroquinone synthase, RifG)	83	100	583	1.46E-211
<b>ctg1_677</b>	ADM46362.1	Nat1 Halogenase	80	99.8	744	3.18E-272
<b>ctg1_678</b>	WP_004559807.1	RifF homolog (Amide synthase)	64	100	344	2.03E-120
<b>ctg1_679</b>	ASZ00151.1	StvE (RifE)	76	101.6	5012	0
<b>ctg1_680</b>	ASZ00150.1	StvD (RifD)	75	100.7	2612	0
<b>ctg1_681</b>	ASZ00149.1	StvC (RifC)	76	99.8	5021	0
<b>ctg1_682</b>	ASZ00148.1	StvB (RifB)	78	101.2	5156	0
<b>ctg1_683</b>	ABV97151.1	RifA	71	105	6781	0
<b>ctg1_684</b>	ASZ00146.1	StvP3 (Rif orf0)	74	100	608	4.03E-220
<b>ctg1_685</b>	ADM46354.1	NADPH dependent azoreductase	56	102.4	238	3.44E-80
<b>ctg1_686</b>	ASZ00140.1	StvO2 (RifT homolog)	61	97.4	345	1.71E-118
<b>ctg1_687</b>	CQR60500.1	CxmS (RifS homolog)	73	97.3	459	2.49E-163
ctg1_688	AWR88388.1	hypothetical_protein	88	99.1	381	3.01E-136
ctg1_689	AWR88387.1	hypothetical_protein	77	99	628	7.26E-228
ctg1_690	AWR88386.1	hypothetical_protein	88	99	328	7.35E-116
ctg1_691	AWR88385.1	putative_GCIN5-related_N-acetyltransferase	73	103.4	1243	0

**Supplementary Figure 4 WAC09654 cluster homology and functional predictions.**

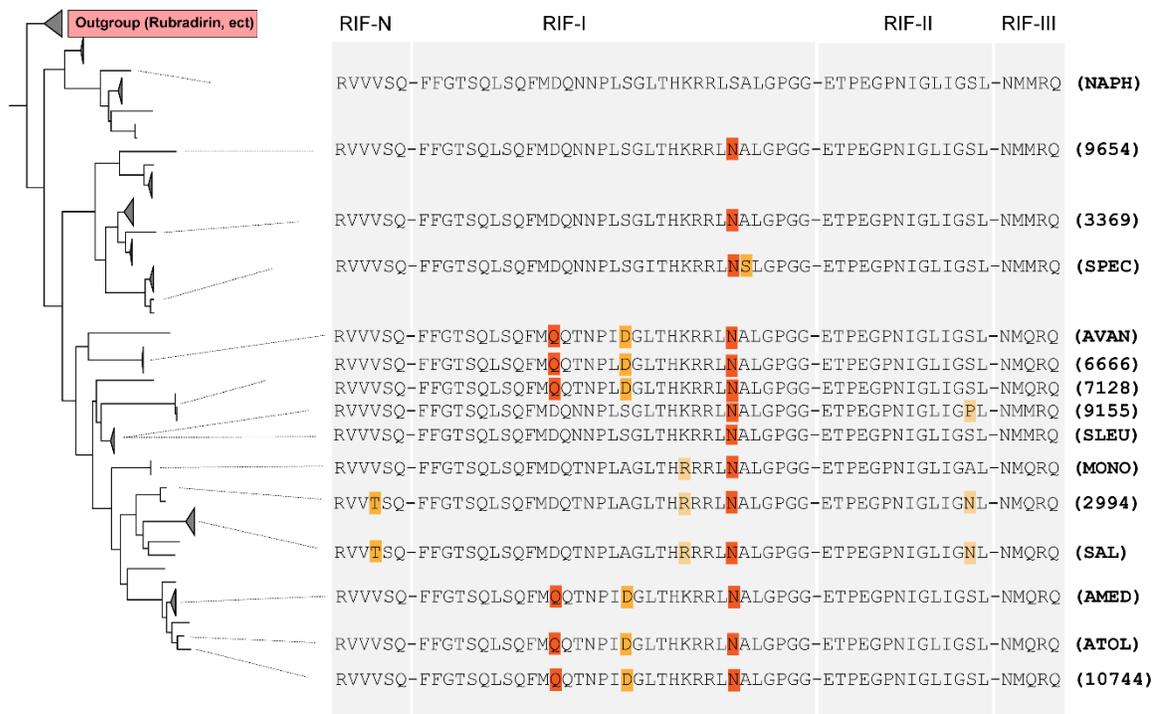
The highest identity match to proteins in the Minimum Information about a Biosynthetic Gene Cluster (MIBiG) database of known clusters<sup>79</sup> for each gene in the WAC09654 locus. Only ctg1\_660 returned no results from this database, and the function was instead inferred from NCBI's conserved domain database. Genes homologous to those in other ansamycins such as Rifamycins (Rif), Streptovaricin (Stv), Naphthomycin (Nat), and Chaxamycin (Cxm), Mitomycin (Mit), and Rubradirin (Rub) were named as they occur in their cluster. Equivalent proteins from the Rifamycin cluster were indicated where possible. Bolded loci represent the predicted bounds of the cluster.



**Supplementary Figure 5 WAC09654 cluster capture.** **A)** Schematic of the region captured. PCR targeted regions labeled cPCR1-3. **B)** Basis of transformation-assisted recombination (TAR) cloning. High molecular weight genomic DNA from WAC09654 was digested with *Psp1406I* and *XbaI*, which cut adjacent to but not within our region of interest, liberating the intact cluster from the genome. This mixture is co-transformed into yeast with linear pCGW containing 50bp homologous regions to each end of the cluster. **C)** Identification of positive clones. Yeast transformants were screened by three PCR reactions which amplify distant regions of the cluster. **D)** Restriction mapping of pCGW:9654. The actual result (left) perfectly matches the theoretical digestions (right). **E)** Sequencing of pCGW:9654. After initial heterologous expression fermentations failed to activate our RNAP inhibition reporter, we sequenced the entire construct using Illumina technology. We were able to assemble the entire 97kbp construct and comparisons to our theoretical construct showed only 6 point mutations (depicted by stars) which were all contained in the plasmid backbone (black) instead of the captured region (blue).



**Supplementary Figure 6 Early rifamycin and WAC09654 biosynthesis.** Synthesis of the naphthoquinone system occurs by a dehydrogenation of Proansamycin X at C8, resulting in the 34a, 8 deoxy rifamycin W intermediate. C8 is then hydroxylated, followed by 34a, which is a prerequisite for oxidative cleavage of rifamycin W and formation of the five-membered ring characteristic of mature, active rifamycins<sup>52</sup>. Compounds putatively produced by *S. coelicolor* M1154 pIJ:luxR pCGW:9654 are highlighted. Halogenation of dideoxy rifamycin W is suggestive of an early chlorination event. Non-halogenated compounds 683, 685, and 707 [M-H]<sup>-</sup> m/z probably represent biosynthesis of mature, rifamycin/streptovaricin like compounds from the dideoxy rifamycin W precursor. Chlorination may be too inefficient to observe the halogenated equivalents of 683/685/707 m/z or alternatively, downstream reactions may be blocked after chlorination.



**Supplementary Figure 7 Putative self-resistance RpoB substitutions in rifamycin-like antibiotic producing organisms.** The sequence of the rifamycin resistance clusters in RpoB (N & RIF-I, II, and III) are given for each producer highlighted in the tree from **Figure 5**. Dark orange residues represent known substitutions that confer moderate to high-level resistance. Orange residues are present in residues that are typically highly conserved and associated with resistance, but the effect of the exact substitution is unknown. Lastly, pale orange represents plausible resistance-conferring substitutions in conserved residues. NAPH – Naphthomycin producer, not *Streptomyces sp.* CS (the whole genome is unavailable for this strain, a closely related producer was used, GCF\_001279545.1). SPEC, *Streptomyces spectabilis* (Streptovaricin, GCF\_007363395.1). AVAN, *Amycolatopsis vancouverensis* (Kanglemycin, GCF\_000388135.1). SLEU, *Streptomyces leeuwenhoekii* (Chaxamycin, GCF\_001013905.1). MONO, *Micromonospora sp.* TP-A0468 (Rifamycin SV, GCF\_020884795.1). SAL, *Salinispora arenicola* (Rifamycin SV, GCF\_000384275.1). AMED, *Amycolatopsis mediterranei* (Rifamycin SV/B, GCF\_000220945.1). ATOL, *Amycolatopsis tolypomycina* (Tolypomycin, GCF\_900105945.1).

## **CHAPTER 6**

### **Discussion**

## SUMMARY

We are entering an era in medicine where the spread of antibiotic-resistant bacteria compromises our ability to manage common infections<sup>29,37,224</sup>. In the face of this challenge, it is more important than ever to understand the full spectrum of biochemical mechanisms employed by bacteria to resist antimicrobial assault, which can help us to discover and develop new therapies. This thesis contains work I performed to advance our understanding of the rifamycin resistome and rifamycin antibiotics, which ultimately stemmed from our attempts to understand the inducible nature of rifamycin resistance in soil Actinobacteria.

I began by building off previous work on inducible rifamycin resistance, which identified a highly conserved *cis*-regulatory element, the RAE, which directs the production of resistance genes in response to rifamycins<sup>213</sup>. In **Chapter 2**, I identified and characterized additional conserved DNA motifs important for induction, which, together with the RAE, form a larger motif resembling a promoter called the extended RAE. I also unequivocally established that induction requires rifamycins to bind and inhibit RNAP. Furthermore, we show that the RAE can discriminate between rifamycin antibiotics and other inhibitors of RNAP. In **Chapter 3**, I used the extended RAE as a beacon to catalog rifamycin resistance genes across the Actinobacteria. We were surprised that in many genera, inducible rifamycin resistance is common. While it is most common in soil saprophytes, we identify this trait in many opportunistic pathogens, such as *Mycobacterium spp.* This analysis also highlighted the abundance and importance of an uncharacterized family of what appeared to be helicase proteins (HelR). **Chapter 4** elucidated the function of HelR and demonstrated that it is not a helicase at all but confers rifamycin resistance by

binding to RNAP and dislodging rifamycin antibiotics. HelRs are the first description of an antibiotic protection protein for RNAP<sup>225</sup>. Finally, in **Chapter 5**, I constructed a cell-based assay for RNAP inhibition based on the RAE and performed a high throughput screen for novel antibiotics. This screen uncovered a family of rifamycins that our data suggest are chlorinated and the first report of a rifamycin that can evade the activity of the rifamycin phosphotransferase Rph.

### **The role of rifamycins and resistance in natural environments**

While the human use of antibiotics has undoubtedly shaped the contemporary environmental resistome, specific mechanisms of rifamycin resistance, the RAE, and the rifamycins themselves are probably millions of years old<sup>53</sup>. From this perspective, it is logical to ask what role antibiotic production and resistance play in natural ecosystems. The biosynthetic loci that encode antibiotics are large (85-120kbp for rifamycins), and the compounds themselves are enormously costly from a metabolic perspective. Rifamycins must fulfill an important role in the ecology of producing organisms to justify this cost. Despite their potent antibacterial activity, the precise function of antibiotics is still hotly debated; some advocate for the view that antibiotics have a function other than growth inhibition in natural environments<sup>76,226,227</sup>. The argument for this perspective goes like this; the concentration of antibiotics produced *in situ* rarely reaches inhibitory concentrations, so how can they ever act as antibiotics? This fact, coupled with observations that sub-inhibitory antibiotics often induce distinct transcriptional responses in target bacteria, has led some to speculate that antibiotics may have a more subtle function than simply

inhibiting surrounding microbes. It is even suggested that antibiotics, specifically rifamycins, may function as signaling molecules, not weapons<sup>228–231</sup>.

While we cannot offer any *in-situ* data, our study of rifamycin resistance is potentially enlightening on the role of rifamycins. Actinobacteria have evolved the perfect system to receive and enact a transcriptional program using a rifamycin signal, the RAE. This *cis*-regulatory mechanism is remarkably specific to rifamycins; despite screening >10 000 extracts and a synthetic library of almost 4000 compounds containing all classes of FDA-approved antibiotics, we have found precisely one non-rifamycin which induces this system, sorangicin. Furthermore, the RAE is also highly sensitive and perfectly capable of detecting sub-inhibitory concentrations relevant to signaling. Conceptually, the RAE would be the ideal way to receive and integrate a rifamycin signal. Instead, bacteria use this mechanism to produce enzymes with one of two functions, to displace rifamycins from RNAP and to destroy them. We don't see any genes with functions that could benefit the rifamycin producer, only mechanisms that promote resistance. Therefore, our data suggest that the role of rifamycins is not subtle, and they do not act as signals. These compounds are chemical weapons designed to restrict the growth of competing microbes.

When discussing this work, I'm often asked, 'Why don't cells just acquire *rpoB* mutations? Why make all these enzymes?'. I think the answer to this question comes down to major differences between how we use rifamycins and how their producers do. Infections regularly contain  $\sim 10^7$ - $10^{10}$  bacteria, a staggeringly large population that will frequently contain at least one spontaneous rifamycin-resistant mutant. It's also roughly equivalent to the average CFU/g of soil. The rifamycin producer *A. mediterranei* is almost certainly not

trying to sterilize an entire gram of soil. Its effects are probably much more local and thus act on far smaller populations of neighboring bacteria. In smaller populations, these mutations will not be nearly as accessible. Furthermore, they may have greater fitness costs in these environments than in the laboratory or during infections. If these mutations are not readily accessible and difficult to maintain due to fitness costs, there is no other option than to evolve specific enzymatic resistance mechanisms. Alternatively, microbes may take measures to suppress the development of these mutations, much like we do, through combination therapy. In **Chapter 5**, we observed the co-production of chaxamycin and chloramphenicol in WAC09155 and rifamycin SV and fidaxomicin in WAC02994. Regardless, the rifamycin resistance substitutions that dominate the resistance landscape in the clinic are not viable strategies for environmental microbes.

Knowledge about antibiotics in their natural environments is extremely limited, and studying microscopic communities directly in the soil is an immense technical challenge. Hopefully, the coming decades will see the development of techniques and methodologies to dissect these important problems. After all, understanding these compounds' chemical ecology may help us find new antibiotics and foresee and prevent the development of resistance.

### **The future of rifamycin antibiotics**

The future will undoubtedly see the deployment of more rifamycin antibiotics in the clinic. Currently, two dual-acting rifamycin-conjugates are in development: a rifamycin-quinolone (TNP-2092) and a rifamycin-nitroimidazole (TNP-2198)<sup>110,112</sup>. These retain activity against cells harboring resistance mutations in one of the two targets, and

present a convenient way to avoid the resistance associated with rifamycin monotherapy. The entry of these dual-acting agents into the market will enable more widespread use of rifamycins against microbes we don't, or rarely, treat with rifamycins currently, such as *Helicobacter pylori*, *C. difficile*, and *Staphylococcus* spp. In the age of resistance, where new antibiotics are hard to come by, this will be good news but also brings the risk of increased selective pressure for resistance. In particular, the effect of rifamycin inactivating enzymes on the activity of the conjugate partner is unknown. As mentioned previously, many Firmicutes possess silent Rph enzymes; using these conjugate molecules *en masse* may select for the overexpression of these determinants.

On a more optimistic note, we are beginning to understand how to evade many forms of rifamycin resistance. In the case of the inactivating enzymes, C25 carbamate derivatives can avoid modification by Arr<sup>203,204</sup>. Our work on Rox revealed that rifamycins locked in the ketone form at C1 intrinsically resist these enzymes<sup>219</sup>. In this work, we have identified a rifamycin that can evade Rph, the subsequent elucidation of its structure will hopefully reveal the rationale for this specific activity and enable the synthesis of Rph-proof compounds. Rgt is the only remaining inactivating enzyme for which no intrinsically resistant rifamycins are known. However, since it modifies the same site as Arr, C25 carbamate derivatives may be similarly poor substrates for Rgt enzymes. Time will tell if researchers can simultaneously incorporate these modifications into a single inactivation-resistant rifamycin. I expect that bypassing the RNAP-protecting activity of HelR enzymes will be significantly more difficult. However, we still lack a detailed structural rationale for the displacement of rifamycins from the binding pocket by HelR. It is also possible that

progress in this area might reveal strategies to bypass this mechanism. The rifamycin resistome is an excellent example of how we can anticipate future resistance threats and proactively work to safeguard antibiotics.

The Achilles heel of rifamycin antibiotics has remained unchanged since their discovery in the 1950s, substitutions in RpoB. Finding new and innovative ways around this problem will be central to the success of rifamycins moving forward. One example is the concept of collateral sensitivity, whereby developing resistance to one antibiotic confers increased sensitivity to another<sup>232</sup>. While combination therapy with rifamycins suppresses the development of resistance by requiring multiple rare mutations to occur at once, compounds with collateral sensitivity benefit from this while also actively selecting against the development of resistance. Discovering compounds that display collateral sensitivity with rifamycins is an attractive area of future research. Along these same lines, rifamycins and related compounds which can bind to resistant alleles of RpoB could have immense utility in rifamycin combination therapy. In fact, such compounds already exist in nature. Kanglemycins, for instance, are extensively decorated rifamycins that make additional contacts on RpoB, allowing them to bind to several rifamycin-resistant alleles<sup>147,171</sup>. Sorangicin binds to the rifamycin binding pocket using a different structural scaffold and retains activity against some common resistant alleles<sup>169,170</sup>. Our high throughput assay for RNAP inhibition by rifamycins could be an excellent way to find compounds that can bind to resistant alleles if we simply perform our assay in *S. venezuelae* with a rifamycin-resistant RNAP. Enzymes and mechanisms associated with inducible

rifamycin resistance are significant problems, but next-generation rifamycins and rifamycin-based therapies should strive to address resistant forms of RNAP.

### **Concluding remarks**

Every cell on earth is, at this very moment, barraged by outside stimuli. Despite this, each cell can process all these stimuli, weigh them against each other, and decide what to do next. Understanding how and why cells make these decisions is the arduous task at the crux of modern biology. All progress made in the realm of rifamycins and the rifamycin resistome during my tenure in the lab stems from the discovery of the RAE. This simple motif unambiguously tells us how cells respond to rifamycin antibiotics, giving us a clear perspective of the often murky and convoluted decision-making process at the cellular level. This enabled us to ask clear and interesting questions across many different contexts. **Chapter 2** poses the question; *how do you know a rifamycin is present?* In **Chapter 3**, we took 3400 genomes and asked *what proteins you make when detecting a rifamycin?* In contrast, during **Chapter 4**, we took a single bacterium and asked *why do you make that specific protein when rifamycins are present?* Finally, in **Chapter 5**, we used that same bacterium to ask: *is there a rifamycin here?*

If there is a singular lesson to be gleaned from this thesis, I believe that identifying and exploiting innate sensing and regulatory mechanisms associated with resistance is the best strategy to characterize antibiotic resistomes writ large. Based on the strength of this approach, one could argue that we know more about the environmental resistome of rifamycins than any other antibiotic.

## REFERENCES

1. Ehrlich, P. Address in Pathology, ON CHEMIOOTHERAPY. *Br. Med. J.* **2**, 359–362 (1913).
2. Williams, K. J. The introduction of ‘chemotherapy’ using arsphenamine - The first magic bullet. *J. R. Soc. Med.* **102**, 343–348 (2009).
3. Riethmiller, S. Elrich, bertheim and atoxyl: The origins of modern chemotherapy. *Bull.Hist.Chem* **23**, 28–33 (1999).
4. Schweitzer, H. EHRLICH’S CHEMOTHERAPY - A NEW SCIENCE. *Science.* **32**, 809–823 (1910).
5. Masters, P. S., O’Bryan, T. A. & Zurlo, J. Trimethoprim-Sulfamethoxazole Revisited. *Arch. Intern. Med.* **163**, 402–410 (2003).
6. Ligon, B. L. Penicillin: Its Discovery and Early Development. *Semin. Pediatr. Infect. Dis.* **15**, 52–57 (2004).
7. Fleming, A. On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
8. Abraham, E. P. *et al.* Further observations on penicillin. *Lancet* **238**, 177–189 (1941).
9. Chain, E. *et al.* Penicillin As a Chemotherapeutic Agent. *Lancet* **236**, 226–228 (1940).
10. Quinn, R. Rethinking antibiotic research and development: World War II and the penicillin collaborative. *Am. J. Public Health* **103**, 426–434 (2013).
11. Brunel, J. Antibiosis from Pasteur to Fleming. *J. Hist. Med. Allied Sci.* **6**, 287–301 (1946).
12. Waksman, S. A. & Woodruff, B. H. The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J. Bacteriol.* **40**, 581–600 (1940).
13. Waksman, S. A. & Woodruff, B. H. Streptothricin, a New Selective Bacteriostatic and Bactericidal Agent, Particularly Active Against Gram-Negative Bacteria. *Proc. Soc. Exp. Biol. Med.* **49**, 207–210 (1942).
14. Waksman, S. A. & Woodruff, B. H. Bacteriostatic and Bactericidal Substances Produced by a Soil Actinomyces. *Proc. Soc. Exp. Biol. Med.* **45**, 609–614 (1940).
15. Waksman, S. A. & Lechevalier, H. A. Neomycin, a new antibiotic active against streptomycin-resistant bacteria, including tuberculosis organisms. *Science (80-. )*. **109**, 305–307 (1949).
16. Schatz, A., Bugie, E. & Waksman, S. A. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Proc. Soc. Exp. Biol. Med.* **55**, 66–69 (1944).
17. Lewis, K. The Science of Antibiotic Discovery. *Cell* **181**, 29–45 (2020).
18. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
19. *Centers for Disease Control and Prevention. Outpatient antibiotic prescriptions - United States, 2021.*
20. Cohen, M. L. Changing patterns of infectious disease. *Nature* **406**, 762–767 (2000).
21. Armstrong, G. L., Conn, L. A. & Pinner, R. W. Trends in infectious disease mortality in the United

- States during the 20th century. *JAMA* **281**, 61–66 (1999).
22. Hughes, V. M. & Datta, N. Conjugative plasmids in bacteria of the ‘pre-antibiotic’ era. *Nature* **302**, 725–726 (1983).
  23. Fleming, A. Penicillin. *Nobel Lect.* 11 (1945).
  24. Barber, M. & Rozwadowska-Dowzenko, M. Infection by penicillin resistant *Staphylococci*. *Lancet* **252**, 641–644 (1948).
  25. O’Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. London: Review on Antimicrobial Resistance, 2016. doi:10.4103/2045-080x.186181
  26. Carlet, J. *et al.* Society’s failure to protect a precious resource: Antibiotics. *Lancet* **378**, 369–371 (2011).
  27. Karakostas, S., Kritsotakis, E. I. & Gikas, A. Pandrug-resistant gram-negative bacteria: A systematic review of current epidemiology, prognosis and treatment options. *J. Antimicrob. Chemother.* **75**, 271–282 (2020).
  28. Magiorakos, A. P. *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: An international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**, 268–281 (2012).
  29. Centers for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States, 2019*. Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2019.
  30. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
  31. Dyar, O. J., Huttner, B., Schouten, J. & Pulcini, C. What is antimicrobial stewardship? *Clin. Microbiol. Infect.* **23**, 793–798 (2017).
  32. Fleming-Dutra, K. E. *et al.* Prevalence of inappropriate antibiotic prescriptions among us ambulatory care visits, 2010–2011. *JAMA - J. Am. Med. Assoc.* **315**, 1864–1873 (2016).
  33. Schwartz, K. L. *et al.* Unnecessary antibiotic prescribing in a Canadian primary care setting: a descriptive analysis using routinely collected electronic medical record data. *C. open* **8**, E360–E369 (2020).
  34. Mulchandani, R., Yang, Y., Gilbert, M. & Van Boeckel, T. Global trends in antimicrobial use in food-producing animals: 2020 to 2030. *PLOS Glob Public Heal.* **3**, 1–11 (2023).
  35. Wolfgang, W. Medical Consequences of Antibiotic Use in Agriculture. *Science (80- )*. **279**, 996–997 (1998).
  36. Årdal, C. *et al.* Antibiotic development — economic, regulatory and societal challenges. *Nat. Rev. Microbiol.* **18**, 267–274 (2020).
  37. Cook, M. A. & Wright, G. D. The past, present, and future of antibiotics. *Sci. Transl. Med.* **14**, (2022).
  38. Wright, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* **5**, 175–186 (2007).
  39. Goldstein, B. P. Resistance to rifampicin: a review. *J. Antibiot. (Tokyo)*. **67**, 625–630 (2014).
  40. Yoshida, H., Bogaki, M., Nakamura, M., Yamanaka, L. M. & Nakamura, S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. *Antimicrob. Agents*

- Chemother.* **34**, 1271–1272 (1990).
41. Maskell, J. P., Sefton, A. M. & Hall, L. M. C. Multiple mutations modulate the function of dihydrofolate reductase in trimethoprim-resistant *Streptococcus pneumoniae*. *Antimicrob. Agents Chemother.* **45**, 1104–1108 (2001).
  42. Sandegren, L. & Andersson, D. I. Bacterial gene amplification: Implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.* **7**, 578–588 (2009).
  43. Gunn, J. S. The *Salmonella* PmrAB regulon: lipopolysaccharide modifications, antimicrobial peptide resistance and more. *Trends Microbiol.* **16**, 284–290 (2008).
  44. Liu, Y. Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: A microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).
  45. Zarubica, T., Baker, M. R., Wright, H. T. & Rife, J. P. The aminoglycoside resistance methyltransferases from the ArmA/Rmt family operate late in the 30S ribosomal biogenesis pathway. *RNA* **17**, 346–355 (2011).
  46. Tran, J. H., Jacoby, G. A. & Hooper, D. C. Interaction of the plasmid-encoded quinolone resistance protein QnrA with *Escherichia coli* topoisomerase IV. *Antimicrob. Agents Chemother.* **49**, 3050–3052 (2005).
  47. Dönhöfer, A. *et al.* Structural basis for TetM-mediated tetracycline resistance. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16900–16905 (2012).
  48. Su, W. *et al.* Ribosome protection by antibiotic resistance ATP-binding cassette protein. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 5157–5162 (2018).
  49. Chesneau, O., Tsvetkova, K. & Courvalin, P. Resistance phenotypes conferred by macrolide phosphotransferases. *FEMS Microbiol. Lett.* **269**, 317–322 (2007).
  50. Webber, M. A. & Piddock, L. J. V. The importance of efflux pumps in bacterial antibiotic resistance. *J. Antimicrob. Chemother.* **51**, 9–11 (2003).
  51. Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).
  52. Bonomo, R. A. & Szabo, D. Mechanisms of multidrug resistance in *Acinetobacter* species and *Pseudomonas aeruginosa*. *Clin. Infect. Dis.* **43**, 49–56 (2006).
  53. Waglechner, N., Elizabeth, C. J. & Wright, G. D. Ancient Antibiotics, Ancient Resistance. *EcoSal Plus* **9**, eESP-0027-02020 (2021).
  54. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).
  55. Bennett, P. M. Plasmid encoded antibiotic resistance: Acquisition and transfer of antibiotic resistance genes in bacteria. *Br. J. Pharmacol.* **153**, 347–357 (2008).
  56. MacLean, R. C. & Millan, A. S. The evolution of antibiotic resistance. *Science.* **365**, 1082–1083 (2019).
  57. Arnold, B. J., Huang, I. T. & Hanage, W. P. Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* **20**, 206–218 (2022).
  58. Kopotsa, K., Osei Sekyere, J. & Mbelle, N. M. Plasmid evolution in carbapenemase-producing

- Enterobacteriaceae: a review. *Ann. N. Y. Acad. Sci.* **1457**, 61–91 (2019).
59. Carattoli, A., Villa, L., Fortini, D. & García-Fernández, A. Contemporary IncII plasmids involved in the transmission and spread of antimicrobial resistance in Enterobacteriaceae. *Plasmid* **118**, (2021).
  60. Abraham, E. P. & Chain, E. An Enzyme from Bacteria able to Destroy Penicillin. *Nature* **146**, 837 (1940).
  61. Wright, G. D. Antibiotic resistance in the environment: A link to the clinic? *Curr. Opin. Microbiol.* **13**, 589–594 (2010).
  62. Courvalin, P. Vancomycin resistance in gram-positive cocci. *Clin. Infect. Dis.* **42**, 25–34 (2006).
  63. Guardabassi, L., Perichon, B., Van Heijenoort, J., Blanot, D. & Courvalin, P. Glycopeptide resistance vanA operons in *Paenibacillus* strains isolated from soil. *Antimicrob. Agents Chemother.* **49**, 4227–4233 (2005).
  64. Rossolini, G. M., D’Andrea, M. M. & Mugnaioli, C. The spread of CTX-M-type extended-spectrum  $\beta$ -lactamases. *Clin. Microbiol. Infect.* **14**, 33–41 (2008).
  65. Barlow, M. *et al.* High rate of mobilization for blaCTX-MS. *Emerg. Infect. Dis.* **14**, 423–428 (2008).
  66. Emerson, R. L., Whiffen, A. J., Bohonos, N. & Deboer, C. Studies on the production of antibiotics by actinomycetes and molds. *J. Bacteriol.* **52**, 357–366 (1946).
  67. Chevrette, M. G. *et al.* The antimicrobial potential of *Streptomyces* from insect microbiomes. *Nat. Commun.* **10**, 1–11 (2019).
  68. D’Costa, V. M., McGrann, K., Hughes, D. W. & Wright, G. D. Sampling the Antibiotic Resistome. *Science* **311**, 374–378 (2006).
  69. Allen, H. K., Moe, L. a., Rodbumrer, J., Gaarder, A. & Handelsman, J. Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J.* **3**, 243–251 (2009).
  70. Dantas, G., Sommer, M. O. A., Oluwasegun, R. & Church, G. M. Bacteria Subsisting on Antibiotics. *Science* **320**, 100–103 (2008).
  71. Forsberg, K. J. *et al.* The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science* **337**, 1107–1111 (2012).
  72. Forsberg, K. J. *et al.* Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**, 612–616 (2014).
  73. Riesenfeld, C. S., Goodman, R. M. & Handelsman, J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* **6**, 981–989 (2004).
  74. Walsh, F. & Duffy, B. The Culturable Soil Antibiotic Resistome: A Community of Multi-Drug Resistant Bacteria. *PLoS One* **8**, (2013).
  75. Bhullar, K. *et al.* Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7**, 1–11 (2012).
  76. Davies, J., Spiegelman, G. B. & Yim, G. The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* **9**, 445–453 (2006).
  77. Pawlowski, A. C. *et al.* A diverse intrinsic antibiotic resistome from a cave bacterium. *Nat. Commun.* **7**, 1–10 (2016).

78. D’Costa, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477**, 457–461 (2011).
79. Waglechner, N., McArthur, A. G. & Wright, G. D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* **4**, 1862–1871 (2019).
80. Hall, B. G. & Barlow, M. Evolution of the serine  $\beta$ -lactamases: Past, present and future. *Drug Resist. Updat.* **7**, 111–123 (2004).
81. Lechevalier, M. P., Prauser, H., Labeda, D. P. & Ruan, J.-S. Two New Genera of Nocardioform Actinomycetes: *Amycolata* gen. nov. and *Amycolatopsis* gen. nov. *Int. J. Syst. Bacteriol.* **36**, 29–37 (1986).
82. Sensi, P. History of the Development of Rifampin. *Rev. Infect. Dis.* **5**, S402–S406 (1983).
83. Sensi, P., Timbal, M. T. & Maffh, G. Rifomycin IX two new antibiotics of rifomycin family: Rifomycin S and rifomycin SV - Preliminary report. *Experientia* **16**, 412 (1960).
84. Maggi, N., Pasqualucci, C. R., Ballotta, R. & Sensi, P. Rifampicin: A New Orally Active Rifamycin. *Chemotherapy (Basel)*. **11**, 285–292 (1966).
85. WHO. *Global Tuberculosis Report 2022*. World Health Organization; 2022.
86. Barberis, I., Bragazzi, N. L., Galluzzo, L. & Martini, M. The history of tuberculosis: From the first historical records to the isolation of Koch’s bacillus. *J. Prev. Med. Hyg.* **58**, E9–E12 (2017).
87. Glaziou, P., Sismanidis, C., Floyd, K. & Raviglione, M. Global epidemiology of tuberculosis. *Cold Spring Harb. Perspect. Med.* **5**, 1–17 (2015).
88. Daffé, M. & Marrakchi, H. Unraveling the structure of the mycobacterial envelope. *Gram-Positive Pathog.* 1087–1095 (2019). doi:10.1128/9781683670131.ch65
89. Gengenbacher, M. & Kaufmann, S. H. E. *Mycobacterium tuberculosis*: Success through dormancy. *FEMS Microbiol. Rev.* **36**, 514–532 (2012).
90. Pai, M. *et al.* Tuberculosis. *Nat. Rev. Dis. Prim.* **2**, (2016).
91. Murray, J. F., Schraufnagel, D. E. & Hopewell, P. C. Treatment of tuberculosis: A historical perspective. *Ann. Am. Thorac. Soc.* **12**, 1749–1759 (2015).
92. Fox, W., Ellard, G. A. & Mitchison, D. A. Studies on the treatment of tuberculosis undertaken by the British Medical Research Council tuberculosis units, 1946-1986, with relevant subsequent publications. *Int. J. Tuberculosis Lung Dis.* **3**, 231–279 (1999).
93. Association, B. T. and T. Short course chemotherapy for pulmonary tuberculosis. A Controlled Trial by the British Thoracic and Tuberculosis Association. *Lancet* **2**, 1102–1104 (1976).
94. Singapore Tuberculosis Service/British Medical Research Council. Clinical trial of six-month and four-month regimens of chemotherapy in the treatment of pulmonary tuberculosis: The results up to 30 months. *Am. Rev. Respir. Dis.* **119**, 579–585 (1979).
95. Johnston, J. C., Chiang, L. & Elwood, K. *Mycobacterium kansasii*. *Microbiol. Spectr.* **5**, 1–9 (2017).
96. WHO. Towards zero leprosy Global Leprosy (Hansen’s disease) Strategy 2021-2030. WHO; 2021.
97. Rothstein, D. M. Rifamycins, alone and in combination. *Cold Spring Harb. Perspect. Med.* **6**, 1–19 (2016).
98. Beldman, M. *et al.* If, when, and how to use rifampin in acute staphylococcal periprosthetic joint

- infections, a multicentre observational study. *Clin. Infect. Dis.* **73**, 1634–1641 (2021).
99. Karlsen, Ø. E. *et al.* Rifampin combination therapy in staphylococcal prosthetic joint infections: a randomized controlled trial. *J. Orthop. Surg. Res.* **15**, 1–9 (2020).
  100. Bassetti, M., Castaldo, N., Cadeo, B. & Cernelutti, A. Prosthetic joint infections: Clinical management, diagnosis, and treatment. *Curr. Opin. Infect. Dis.* **32**, 102–112 (2019).
  101. Thwaites, G. E. *et al.* Adjunctive rifampicin for *Staphylococcus aureus* bacteraemia (ARREST): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet* **391**, 668–678 (2018).
  102. Deal, W. B. & Sanders, E. Efficacy of Rifampin in Treatment of Meningococcal Carriers. *N. Engl. J. Med.* **281**, 641–645 (1969).
  103. Masr, H. *et al.* Recommendations on prophylaxis and therapy for disseminated *Mycobacterium avium* complex disease in patients infected with the human immunodeficiency virus. *N. Engl. J. Med.* **329**, 898–904 (1993).
  104. Baciewicz, A. M. & Self, T. H. Rifampin Drug Interactions. *Arch. Intern. Med.* **144**, 1667–1671 (1984).
  105. Blaschke, T. F. & Skinner, Mi. H. The Clinical Pharmacokinetics of Rifabutin. *Clin. Infect. Dis.* **22**, 15–22 (1996).
  106. Iseman, M. D. Tuberculosis therapy: Past, present and future. *Eur. Respir. Journal, Suppl.* **20**, 87–94 (2002).
  107. Dorman, S. E. *et al.* Four-Month Rifapentine Regimens with or without Moxifloxacin for Tuberculosis. *N. Engl. J. Med.* **384**, 1705–1718 (2021).
  108. Koo, H. L. & DuPont, H. L. Rifaximin: A Unique Gastrointestinal-Selective Antibiotic for Enteric Diseases. *Curr. Opin. Gastroenterol.* **23**, 17–25 (2010).
  109. Pimentel, M. *et al.* Rifaximin Therapy for Patients with Irritable Bowel Syndrome without Constipation. *N. Engl. J. Med.* **364**, 22–32 (2011).
  110. Ma, Z. & Lynch, A. S. Development of a Dual-Acting Antibacterial Agent (TNP-2092) for the Treatment of Persistent Bacterial Infections. *J. Med. Chem.* **59**, 6645–6657 (2016).
  111. Butler, M. S. *et al.* Analysis of the Clinical Pipeline of Treatments for Drug-Resistant Bacterial Infections: Despite Progress, More Action Is Needed. *Antimicrob. Agents Chemother.* **66**, 1–20 (2022).
  112. Ma, Z. *et al.* Design, Synthesis, and Characterization of TNP-2198, a Dual-Targeted Rifamycin-Nitroimidazole Conjugate with Potent Activity against Microaerophilic and Anaerobic Bacterial Pathogens. *J. Med. Chem.* **65**, 4481–4495 (2022).
  113. Ang, C. W., Jarrad, A. M., Cooper, M. A. & Blaskovich, M. A. T. Nitroimidazoles: Molecular Fireworks That Combat a Broad Spectrum of Infectious Diseases. *J. Med. Chem.* **60**, 7636–7657 (2017).
  114. Molodtsov, V. *et al.* X-ray crystal structures of the *Escherichia coli* RNA polymerase in complex with Benzoxazinorifamycins. *J. Med. Chem.* **56**, 4758–4763 (2013).
  115. Rothstein, D. M., Van Duzer, J., Sternlicht, A. & Gilman, S. C. Rifalazil and other benzoxazinorifamycins in the treatment of chlamydia-based persistent infections. *Arch. Pharm. (Weinheim)*. **340**, 517–529 (2007).

116. Rajeswaran, W. *et al.* Optimization of Benzoxazinorifamycins to Improve *Mycobacterium tuberculosis* RNA Polymerase Inhibition and Treatment of Tuberculosis. *ACS Infect. Dis.* **8**, 1422–1438 (2022).
117. Ashkar, S. R. *et al.* Optimization of Benzoxazinorifamycins to Minimize hPXR Activation for the Treatment of Tuberculosis and HIV Coinfection. *ACS Infect. Dis.* **8**, 1408–1421 (2022).
118. August, P. R. *et al.* Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**, 69–79 (1998).
119. Floss, H. G. & Yu, T.-W. Rifamycin Mode of Action, Resistance, and Biosynthesis. *Chem. rev.* **105**, 621–632 (2005).
120. Floss, H. G., Yu, T. W. & Arakawa, K. The biosynthesis of 3-amino-5-hydroxybenzoic acid (AHBA), the precursor of mC 7 N units in ansamycin and mitomycin antibiotics: A review. *J. Antibiot. (Tokyo)*. **64**, 35–44 (2011).
121. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496 (2006).
122. Xu, J., Wan, E., Kim, C. J., Floss, H. G. & Mahmud, T. Identification of tailoring genes involved in the modification of the polyketide backbone of rifamycin B by *Amycolatopsis mediterranei* S699. *Microbiology* **151**, 2515–2528 (2005).
123. Wang, N.-J. *et al.* Isolation and structure of a new ansamycin antibiotic kanglemycin A from a *Nocardia*. *J. Antibiot. (Tokyo)*. **41**, 264–267 (1988).
124. Xiao, Y. S. *et al.* Rifamorpholines A–E, potential antibiotics from locust-associated actinobacteria *Amycolatopsis* sp. Hca4. *Org. Biomol. Chem.* 3909–3916 (2017). doi:10.1039/c7ob00614d
125. Williams, D. E. *et al.* Aminorifamycins and Sporalactams Produced in Culture by a *Micromonospora* sp. Isolated from a Northeastern-Pacific Marine Sediment Are Potent Antibiotics. *Org. Lett.* **19**, 766–769 (2017).
126. Rinehart, K. L. *et al.* Chemistry of the Streptovaricins. VIII. Structures of Streptovaricin A, B, D, E, F, and G. *J. Am.* **93**, 6273–6274 (1971).
127. Ganguly, A. K., Liu, Y. T., Sarre, O. Z. & Szmulewicz, S. Structures of Halomicins A and C. *J. Antibiot. (Tokyo)*. **30**, 625–627 (1977).
128. Kishi, T., Harada, S., Asai, M., Muroi, M. & Mizuno, K. Tolypomycin. II. Structures of Tolyposamine and Tolypomycin Y. *Tetrahedron Lett.* **2**, 97–100 (1969).
129. Rateb, M. E. *et al.* Chaxamycins A - D, bioactive ansamycins from a hyper-arid desert *Streptomyces* sp. *J. Nat. Prod.* **74**, 1491–1499 (2011).
130. Madigan, M. *et al.* *Brock Biology of Microorganisms, 14th edition.* (Pearson, 2014).
131. Chen, J., Boyaci, H. & Campbell, E. A. Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **19**, 95–109 (2021).
132. Vassilyev, D. G., Vassilyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157–162 (2007).
133. Li, L., Molodtsov, V., Lin, W., Ebright, R. H. & Zhang, Y. RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. *Proc. Natl. Acad. Sci.* **117**, 5801–5809 (2020).

134. Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460–464 (2005).
135. Chen, J., Boyaci, H. & Campbell, E. A. Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **19**, 95–109 (2021).
136. You, L. *et al.* Structural basis for intrinsic transcription termination. *Nature* **613**, 783–789 (2023).
137. Molodtsov, V., Wang, C., Firlar, E., Kaelber, J. T. & Ebright, R. H. Structural basis of Rho-dependent transcription termination. *Nature* **614**, 367–374 (2023).
138. Kirsch, S. H., Haeckl, F. P. J. & Müller, R. Beyond the approved : target sites and inhibitors of bacterial RNA polymerase from bacteria and fungi. *Nat. Prod. Rep.* **39**, 1226–1263 (2022).
139. Calvori, C., Frontali, L., Leoni, L. & Tecce, G. Effect of Rifamycin on Protein Synthesis. *Nature* **207**, 417–418 (1965).
140. Sippel, A. & Hartmann, G. Mode of action of rifamycin on the RNA polymerase reaction. *Biochim. Biophys. Acta* **157**, 218–219 (1968).
141. Hartmann, G., Honikel, K. O., Knusel, F. & Nuesch, J. The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochim. Biophys. Acta* **145**, 843–844 (1967).
142. Wehrli, W., Knusel, F. & Staehelin, M. Action of Rifamycin on RNA-Polymerase from Sensitive and Resistant Bacteria. *Biochem. Biophys. Res. Commun.* **32**, 284–288 (1968).
143. McClure, W. R. & Cech, C. L. On the mechanism of rifampicin inhibition of RNA synthesis. *J. Biol. Chem.* **253**, 8949–8956 (1978).
144. Campbell, E. A. *et al.* Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **104**, 901–912 (2001).
145. Lin, W. *et al.* Structural Basis of *Mycobacterium tuberculosis* Transcription and Transcription Inhibition. *Mol. Cell* **66**, 169–179.e8 (2017).
146. Feklistov, A. *et al.* Rifamycins do not function by allosteric modulation of binding of Mg<sup>2+</sup> to the RNA polymerase active center. *Proc. Natl. Acad. Sci.* **105**, 14820–14825 (2008).
147. Peek, J. *et al.* Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nat. Commun.* **9**, (2018).
148. Molodtsov, V., Scharf, N. T., Stefan, M. A., Garcia, G. A. & Murakami, K. S. Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **103**, 1034–1045 (2017).
149. Campbell, E. A. *et al.* Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **104**, 901–912 (2001).
150. Farrell, D. J. *et al.* In Vitro activity and single-step mutational analysis of rifamycin SV tested against enteropathogens associated with traveler’s diarrhea and *Clostridium difficile*. *Antimicrob. Agents Chemother.* **55**, 992–996 (2011).
151. O’Neill, A. J., Cove, J. H. & Chopra, I. Mutation frequencies for resistance to fusidic acid and rifampicin in *Staphylococcus aureus*. *J. Antimicrob. Chemother.* **47**, 647–650 (2001).
152. David, H. L. Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*. *Appl. Microbiol.* **20**, 810–814 (1970).
153. Nyinoh, I. W. Spontaneous mutations conferring antibiotic resistance to antitubercular drugs at a

- range of concentrations in *Mycobacterium smegmatis*. *Drug Dev. Res.* **80**, 147–154 (2019).
154. Bush, K. & Bradford, P. A.  $\beta$ -Lactams and  $\beta$ -Lactamase Inhibitors An Overview. *Cold Spring Harb. Perspect. Medicine* **6**, 1–22 (2016).
  155. Fàbrega, A., Madurga, S., Giralt, E. & Vila, J. Mechanism of action of and resistance to quinolones. *Microb. Biotechnol.* **2**, 40–61 (2009).
  156. Lin, J., Zhou, D., Steitz, T. A., Polikanov, Y. S. & Gagnon, M. G. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annu. Rev. Biochem.* **87**, 451–478 (2018).
  157. Ohno, H., Koga, H., Kohno, S., Tashiro, T. & Hara, K. Relationship between rifampin MICs for and rpoB mutations of *Mycobacterium tuberculosis* strains isolated in Japan. *Antimicrob. Agents Chemother.* **40**, 1053–1056 (1996).
  158. Kurepina, N., Chudaev, M., Kreiswirth, B. N., Nikiforov, V. & Mustaev, A. Mutations compensating for the fitness cost of rifampicin resistance in *Escherichia coli* exert pleiotropic effect on RNA polymerase catalysis. *Nucleic Acids Res.* **50**, 5739–5756 (2022).
  159. Jin, D. J. & Gross, C. A. Characterization of the pleiotropic phenotypes of rifampin-resistant rpoB mutants of *Escherichia coli*. *J. Bacteriol.* **171**, 5229–5231 (1989).
  160. Hu, H. & Ochi, K. Novel Approach for Improving the Productivity of Antibiotic-Producing Strains by Inducing Combined Resistant Mutations. *Appl. Environ. Microbiol.* **67**, 1885–1892 (2001).
  161. Comas, I. *et al.* Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat. Genet.* **44**, 106–110 (2012).
  162. De Vos, M. *et al.* Putative compensatory mutations in the rpoC gene of rifampin-resistant mycobacterium tuberculosis are associated with ongoing transmission. *Antimicrob. Agents Chemother.* **57**, 827–832 (2013).
  163. Heil, A. & Zillig, W. Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett.* **11**, 165–168 (1970).
  164. Jin, D. J. & Gross, C. A. Mapping and sequencing of mutations in the *Escherichia coli* rpoB gene that lead to rifampicin resistance. *J. Mol. Biol.* **202**, 45–58 (1988).
  165. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber. Lung Dis.* **79**, 3–29 (1998).
  166. Ovchinnikov, Y. A. *et al.* RNA polymerase rifampicin resistance mutations in *Escherichia coli*: Sequence changes and dominance. *Mol. Gen. Genet.* **190**, 344–348 (1983).
  167. Kim, H. *et al.* Mechanism of natural rifampin resistance of *Streptomyces* spp. *Syst. Appl. Microbiol.* **28**, 398–404 (2005).
  168. Ishikawa, J., Chiba, K., Kurita, H. & Satoh, H. Contribution of rpoB2 RNA polymerase beta subunit gene to rifampin resistance in *Nocardia* species. *Antimicrob. Agents Chemother.* **50**, 1342–1346 (2006).
  169. Campbell, E. A. *et al.* Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674–682 (2005).
  170. Lilic, M. *et al.* The antibiotic sorangicin A inhibits promoter DNA unwinding in a *Mycobacterium*

- tuberculosis* rifampicin-resistant RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 30423–30432 (2020).
171. Mosaei, H. *et al.* Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant *Mycobacterium tuberculosis*. *Mol. Cell* **72**, 263–274.e5 (2018).
  172. Peek, J. *et al.* A Semisynthetic Kanglemycin Shows in Vivo Efficacy against High-Burden Rifampicin Resistant Pathogens. *ACS Infect. Dis.* **6**, 2431–2440 (2020).
  173. King, A. M. *et al.* Aspergillomarasmine A overcomes metallo beta-lactamase antibiotic resistance. *Nature* **510**, 503–506 (2014).
  174. Balibar, C. J. & Grabowicz, M. Mutant alleles of lptD increase the permeability of *Pseudomonas aeruginosa* and define determinants of intrinsic resistance to antibiotics. *Antimicrob. Agents Chemother.* **60**, 845–854 (2016).
  175. French, S. *et al.* Potentiation of Antibiotics against Gram-Negative Bacteria by Polymyxin B Analogue SPR741 from Unique Perturbation of the Outer Membrane. *ACS Infect. Dis.* **6**, 1405–1412 (2020).
  176. Stokes, J. M. *et al.* Pentamidine sensitizes Gram-negative pathogens to antibiotics and overcomes acquired colistin resistance. *Nat. Microbiol.* **2**, (2017).
  177. Venter, H., Mowla, R., Ohene-Agyei, T. & Ma, S. RND-type drug efflux pumps from Gram-negative bacteria: Molecular mechanism and inhibition. *Front. Microbiol.* **6**, 1–11 (2015).
  178. Pang, Y. *et al.* Study of the rifampin monoresistance mechanism in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **57**, 893–900 (2013).
  179. Li, G. *et al.* Study of efflux pump gene expression in rifampicin-monoresistant *Mycobacterium tuberculosis* clinical isolates. *J. Antibiot. (Tokyo)*. **68**, 431–435 (2015).
  180. Ramón-García, S., Martín, C., Thompson, C. J. & Ainsa, J. A. Role of the *Mycobacterium tuberculosis* P55 efflux pump in intrinsic drug resistance, oxidative stress responses, and growth. *Antimicrob. Agents Chemother.* **53**, 3675–3682 (2009).
  181. Hurst-Hess, K. *et al.* Mycobacterial SigA and SigB cotranscribe essential housekeeping genes during exponential growth. *MBio* **10**, 1–17 (2019).
  182. Newell, K. V, Thomas, D. P., Brekasis, D. & Paget, M. S. B. The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on *Streptomyces coelicolor*. *Mol. Microbiol.* **60**, 687–696 (2006).
  183. Verma, A. K. & Chatterji, D. Dual role of MsRbpA: Transcription activation and rescue of transcription from the inhibitory effect of rifampicin. *Microbiol. (United Kingdom)* **160**, 2018–2029 (2014).
  184. Hubin, E. A. *et al.* Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *Elife* **6**, 1–40 (2017).
  185. Bortoluzzi, A. *et al.* *Mycobacterium tuberculosis* RNA polymerase-binding protein A (RbpA) and its interactions with sigma factors. *J. Biol. Chem.* **288**, 14438–14450 (2013).
  186. Hu, Y., Morichaud, Z., Chen, S., Leonetti, J. P. & Brodolin, K. *Mycobacterium tuberculosis* RbpA protein is a new type of transcriptional activator that stabilizes the  $\sigma$  a-containing RNA polymerase holoenzyme. *Nucleic Acids Res.* **40**, 6547–6557 (2012).
  187. Surette, M. D., Spanogiannopoulos, P. & Wright, G. D. The Enzymes of the Rifamycin Antibiotic

- Resistome. *Acc. Chem. Res.* **111**, (2021).
188. Dabbs, E. R. *et al.* Ribosylation by mycobacterial strains as a new mechanism of rifampin inactivation. *Antimicrob. Agents Chemother.* **39**, 1007–1009 (1995).
  189. Tanaka, Y. *et al.* Different rifampicin inactivation mechanisms in *Nocardia* and related taxa. *Microbiol. Immunol.* **40**, 1–4 (1996).
  190. Quan, S., Venter, H. & Dabbs, E. R. Ribosylative Inactivation of Rifampin by *Mycobacterium smegmatis* Is a Principal Contributor to Its Low Susceptibility to This Antibiotic. *Antimicrob. Agents Chemother.* **41**, 2456–2460 (1997).
  191. Quan, S. *et al.* ADP-ribosylation as an intermediate step in inactivation of rifampin by a mycobacterial gene. *Antimicrob. Agents Chemother.* **43**, 181–184 (1999).
  192. Baysarowich, J. *et al.* Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 4886–4891 (2008).
  193. Cohen, M. S. & Chang, P. Insights into the biogenesis, function, and regulation of ADP-ribosylation. *Nat. Chem. Biol.* **14**, 236–243 (2018).
  194. Schuller, M. *et al.* Molecular basis for DarT ADP-ribosylation of a DNA base. *Nature* **596**, 597–602 (2021).
  195. Bullen, N. P. *et al.* An ADP-ribosyltransferase toxin kills bacterial cells by modifying structured non-coding RNAs. *Mol. Cell* **82**, 3484-3498.e11 (2022).
  196. Victoria, L., Gupta, A., Gómez, J. L. & Robledo, J. *Mycobacterium abscessus* complex: A Review of Recent Developments in an Emerging Pathogen. *Front. Cell. Infect. Microbiol.* **11**, 1–8 (2021).
  197. Floto, R. A. *et al.* US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis. *Thorax* **71**, i1–i22 (2016).
  198. Rominski, A., Roditscheff, A., Selchow, P., Böttger, E. C. & Sander, P. Intrinsic rifamycin resistance of *Mycobacterium abscessus* is mediated by ADP-ribosyltransferase MAB\_0591. *J. Antimicrob. Chemother.* dkw466 (2016). doi:10.1093/jac/dkw466
  199. Tribuddharat, C. & Fennewald, M. Integron-mediated rifampin resistance in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **43**, 960–962 (1999).
  200. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **51**, D690–D699 (2023).
  201. CARD prevalence data; Arr-2. (2023). Available at: <https://card.mcmaster.ca/ontology/39281>. (Accessed: 22nd February 2023)
  202. Steffen, R. *et al.* Rifamycin SV-MMX® for treatment of travellers' diarrhea: Equally effective as ciprofloxacin and not associated with the acquisition of multi-drug resistant bacteria. *J. Travel Med.* **25**, 1–11 (2018).
  203. Paulowski, L. *et al.* C25-modified rifamycin derivatives with improved activity against *Mycobacterium abscessus*. *PNAS Nexus* **1**, 1–13 (2022).
  204. Combrink, K. D. *et al.* New C25 carbamate rifamycin derivatives are resistant to inactivation by ADP-ribosyl transferases. *Bioorg. Med. Chem. Lett.* **17**, 522–526 (2007).

205. Lan, T. *et al.* Redesign of Rifamycin Antibiotics to Overcome ADP-Ribosylation-Mediated Resistance. *Angew. Chemie* **134**, (2022).
206. Harbottle, J., Mosaei, H., Allenby, N. & Zenkin, N. Kanglemycin a can overcome rifamycin resistance caused by adp-ribosylation by arr protein. *Antimicrob. Agents Chemother.* **65**, (2021).
207. Yazawa, K. *et al.* Inactivation of rifampin by *Nocardia brasiliensis*. *Antimicrob. Agents Chemother.* **37**, 1313–1317 (1993).
208. Spanogiannopoulos, P. Exploring Rifamycin Inactivation from the Soil Microbiome. (McMaster, 2014).
209. Spanogiannopoulos, P., Thaker, M., Koteva, K., Waglechner, N. & Wright, G. D. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob. Agents Chemother.* **56**, 5061–5069 (2012).
210. Shida, O., Takagi, H., Kadowaki, K. & Miyaji, M. Rifampicin Inactivation by *Bacillus* species. *J. Antibiot. (Tokyo)*. **48**, 815–819 (1995).
211. Dabbs, E. Rifampicin inactivation by *Rhodococcus* and *Mycobacterium* species. *FEMS Microbiol. Lett.* **44**, 395–399 (1987).
212. Yazawa, K., Mikami, Y., Maeda, A. & Morisaki, N. Phosphorylative inactivation of rifampicin by *Nocardia otitidiscaviarum*. *J. Antimicrob. Chemother.* **33**, 1127–1135 (1994).
213. Spanogiannopoulos, P., Waglechner, N., Koteva, K. & Wright, G. D. A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7102–7107 (2014).
214. Stogios, P. J. *et al.* Rifampin phosphotransferase is an unusual antibiotic resistance kinase. *Nat. Commun.* **7**, (2016).
215. Qi, X. *et al.* Structural basis of rifampin inactivation by rifampin phosphotransferase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 3803–3808 (2016).
216. Andersen, S. J., Quan, S., Gowan, B. & Dabbs, E. R. Monooxygenase-like sequence of a *Rhodococcus equi* gene conferring increased resistance to rifampin by inactivating this antibiotic. *Antimicrob. Agents Chemother.* **41**, 218–221 (1997).
217. Mascotti, M. L., Juri Ayub, M., Furnham, N., Thornton, J. M. & Laskowski, R. A. Chopping and Changing: the Evolution of the Flavin-dependent Monooxygenases. *J. Mol. Biol.* **428**, 3131–3146 (2016).
218. Hoshino, Y. *et al.* Monooxygenation of rifampicin catalyzed by the rox gene product of *Nocardia farcinica*: Structure elucidation, gene identification and role in drug resistance. *J. Antibiot. (Tokyo)*. **63**, 23–28 (2010).
219. Koteva, K. *et al.* Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem. Biol.* **25**, 403–412.e5 (2018).
220. Liu, L. K., Dai, Y., Abdelwahab, H., Sobrado, P. & Tanner, J. J. Structural Evidence for Rifampicin Monooxygenase Inactivating Rifampicin by Cleaving Its Ansa-Bridge. *Biochemistry* **57**, 2065–2068 (2018).
221. Zheng, X. F. *et al.* Characterization of the Rifamycin-Degrading Monooxygenase From Rifamycin Producers Implicating Its Involvement in Saliniketal Biosynthesis. *Front. Microbiol.* **11**, 1–8 (2020).

222. Wilson, M. C., Guder, T. A., Mahmud, T. & Moore, B. S. Shared biosynthesis of the saliniketals and rifamycins in *Salinispora arenicola* is controlled by the sare1259 encoded cytochrome P450. *J. Am. Chem. Soc.* **132**, 12757–12765 (2010).
223. Cantón, R. Antibiotic resistance genes from the environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. *Clin. Microbiol. Infect.* **15 Suppl 1**, 20–25 (2009).
224. World Health Organization. *Global antimicrobial resistance and use surveillance system (GLASS) report 2021*. World Health Organization. (2021).
225. Surette, M. D., Waglechner, N., Koteva, K. & Wright, G. D. HelR is a helicase-like protein that protects RNA polymerase from rifamycin antibiotics. *Mol. Cell* **82**, 3151-3165.e9 (2022).
226. Aminov, R. I. The role of antibiotics and antibiotic resistance in nature. *Environ. Microbiol.* **11**, 2970–2988 (2009).
227. Fajardo, A., Linares, J. F. & Martínez, J. L. Towards an ecological approach to antibiotics and antibiotic resistance genes. *Clin. Microbiol. Infect.* **15**, 14–16 (2009).
228. Goh, E.B. *et al.* Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 17025–17030 (2002).
229. Yim, G., De La Cruz, F., Spiegelman, G. B. & Davies, J. Transcription modulation of *Salmonella enterica* serovar typhimurium promoters by sub-MIC levels of rifampin. *J. Bacteriol.* **188**, 7988–7991 (2006).
230. Romero, D., Traxler, M. F., López, D. & Kolter, R. Antibiotics as signal molecules. *Chem. Rev.* **111**, 5492–5505 (2011).
231. Yim, G., Wang, H. H. & Davies, J. Antibiotics as signalling molecules. *Philos. Trans. R. Soc. B Biol. Sci.* **362**, 1195–1200 (2007).
232. Roemhild, R. & Andersson, D. I. Mechanisms and therapeutic potential of collateral sensitivity to antibiotics. *PLoS Pathog.* **17**, 1–6 (2021).