

**Identification and Evaluation of Bacterial RNA Polymerase Inhibitors Using a Novel Plasmid-based Transcription Assay**

**by**

**Nathan Thomas Scharf**

**A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Medicinal Chemistry)  
in the University of Michigan  
2017**

**Doctoral Committee:**

**Professor George A. Garcia, Chair  
Professor David H. Sherman  
Professor Janet L. Smith  
Professor Ronald W. Woodard**

Nathan Thomas Scharf

[scharfn@umich.edu](mailto:scharfn@umich.edu)

ORCID iD: [0000-0002-5867-170X](https://orcid.org/0000-0002-5867-170X)

© Nathan Thomas Scharf 2017

## **Acknowledgements**

I would like to thank my boss George Garcia for his dedication and hard work in the development of not only an independent scientist, but a better person. I would like to thank my parents and siblings for their continued support and encouragement throughout the process. I would also like to thank my fellow Medicinal Chemistry Ph.D. cohort for the comradery and memories over the last five years.

## Table of Contents

<b>Acknowledgements</b>	<b>ii</b>
<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>viii</b>
<b>Abstract</b>	<b>x</b>
<b>Chapter I</b>	
<b>Introduction</b>	<b>1</b>
Tuberculosis and <i>Mycobacterium tuberculosis</i>	2
Rifamycins, Rifampin, and Rifamycin Resistance	5
Known Bacterial RNA Polymerase Inhibitors	8
Malachite Green Aptamer	9
Research Objectives	10
<b>Chapter II</b>	
<b>Development and Optimization of In vitro Plasmid-Based Transcription Assay Utilizing Malachite Green Aptamer Detection</b>	<b>12</b>
Materials and Methods	13
Results	23
Discussion	30
Conclusions	32
Notes to Chapter II	33
<b>Chapter III</b>	
<b>High-throughput Screen of <i>Escherichia coli</i> <math>\beta</math>S531L/<math>\beta</math>'V408G RNA Polymerase against 150,000 Compound Library</b>	<b>34</b>
Materials and Methods	36
Results	45
Discussion	51
Conclusions	57
Notes to Chapter III	59

<b>Chapter IV</b>	
<b>In vitro Evaluation of Hit Scaffolds Identified from High-Throughput Screen</b>	<b>60</b>
Materials and Methods	61
Results	66
Discussion	73
Conclusions	79
Notes to Chapter IV	80
Appendices	81
<b>Chapter V</b>	
<b>In vitro Evaluation of Benzoxazinorifamycin Analogs against Wild-type and Rifamycin-Resistant Mutants of <i>Escherichia coli</i> and <i>Mycobacterium tuberculosis</i> RNA Polymerase</b>	<b>91</b>
Materials and Methods	94
Results	99
Discussion	108
Conclusions	112
Notes to Chapter V	114
Appendices	115
<b>Chapter VI</b>	
<b>Summary, Conclusions, and Future Directions</b>	<b>120</b>

## List of Figures

Figure I-1. Number of new antibiotics approved by year from 1980 – 2017	1
Figure I-2. Structures of rifamycins of particular interest	6
Figure II-1. Overview of modifications to prepare pMGA12 plasmid template	18
Figure II-2. Iterative cloning technique to insert additional MGA repeats	19
Figure II-3. 16% PAGE of synthesized DNA nanocircles	23
Figure II-4. Mass spectra of circularized MG-45	24
Figure II-5. Fluorescence detection of rolling circle transcription and plasmid-based transcription	25
Figure II-6. Plasmid-based transcription assay using different plasmid DNA templates	27
Figure II-7. Inhibition of <i>E. coli</i> WT RNAP by rifampin in plasmid-based transcription assay	29
Figure III-1. Transcription activity of <i>E. coli</i> and MTB Rif <sup>R</sup> RNAP with compensatory mutation	46
Figure III-2. Controls of transcription assay with <i>E. coli</i> $\beta$ S531L/ $\beta$ 'V408G RNAP in 384-well plate format for high-throughput screen	47
Figure III-3. Inhibition plot of 150,554 compounds tested in the HTS against <i>E. coli</i> $\beta$ S531L/ $\beta$ 'V408G RNAP	48
Figure III-4. Flow chart of selection criteria for high-throughput screen	49
Figure III-5. Waterfall plot of IC <sub>50</sub> values determined against <i>E. coli</i> $\beta$ S531L/ $\beta$ 'V408G of 234 compounds identified during HTS triage	50
Figure III-6. Comparison of compound transfer methods for transcription assay inhibition studies	51
Figure IV-1. Covalent interaction of nitrile with L-cysteine to form cysteine adduct	61
Figure IV-2. Plasmid map of DNA template containing terminator sequences and <i>E. coli</i> or MTB promoter	64
Figure IV-3. Overview of transcription assay used to detect reversibility of inhibition	65
Figure IV-4. Fluorescence inhibition of ethidium bromide by compounds that were active in the fluorescent intercalator displacement (FID) assay	66

Figure IV-5. Structures of HTS hits active against MTB in culture and pyrazolopyrimidine analogs synthesized and tested for inhibition of <i>E. coli</i> and MTB RNAP	68
Figure IV-6. Analogs of HTS hit 257178 tested for inhibition against <i>E. coli</i> and MTB WT RNAPs in the plasmid-based transcription assay	70
Figure IV-7. Structure of cyanopyrimidine scaffold tested for activity against <i>E. coli</i> and MTB WT RNAPs	70
Figure IV-8. Plasmid-based transcription assay after RNAP preincubation with compounds	71
Figure IV-9. LC-MS chromatogram of positive control (5-bromopyrimidine-2-carbonitrile) and exemplary cyanopyrimidine (257176)	72
Figure IV-10. Proposed reaction for nitrile reactivity assay	78
Figure IV-A1. Pyrazolopyrimidine analogs tested for inhibition against <i>E. coli</i> and MTB WT RNAPs in the plasmid-based transcription assay	82
Figure IV-A2. Analogs of 23486 tested for inhibition against <i>E. coli</i> and MTB WT RNAPs in plasmid-based transcription assay	83
Figure IV-A3. Analogs of cyanopyrimidine scaffold tested for inhibition against <i>E. coli</i> and MTB WT RNAPs in plasmid-based transcription assay	84
Figure IV-A4. Synthetic scheme for RAR-1-35, RAR-1-54, and intermediates	85
Figure IV-A5. Synthetic scheme for 193157 analogs and intermediates	86
Figure IV-A6. Synthetic scheme for cyanopyrimidines 257176 and 257181	86
Figure IV-A7. LC-MS chromatogram of negative control CCG-257084 and cyanopyrimidines 257180, 257181, and 257182	90
Figure V-1. Structures of benzoxazinorifamycins and the clinical rifamycins rifampin and rifabutin	92
Figure V-2. Rifabutin bound to <i>Thermus thermophilus</i> RNAP (PDB: 2A68)	93
Figure V-3. Schematic of heparin trapping RNAP in an inactive state	98
Figure V-4. X-ray crystal structure of bxRif2c bound to <i>E. coli</i> RNA polymerase	100
Figure V-5. Concentration-response curves of rifampin against <i>E. coli</i> and MTB Rif <sup>R</sup> RNAP using the plasmid-based transcription assay	101
Figure V-6. Time course plasmid-based transcription assay of <i>E. coli</i> and MTB WT RNAP with and without heparin	102
Figure V-7. Interactions of rifampin with residues of <i>E. coli</i> RNAP	103
Figure V-8. Structural bases of Rif <sup>R</sup> by clinically relevant mutations S531L, H526Y, and D516V	104

Figure V-9. Overview of bxRif structure-activity relationship studies against Rif <sup>R</sup> RNAP	107
Figure V-A1. Percent Rif <sup>R</sup> mutant in <i>E. coli</i> and MTB RNAP Purifications	115
Figure V-A2. Electron density maps of RNAP and RNAP•RMP complexes	116



## List of Tables

Table I-1. Drugs used for the treatment of tuberculosis	3
Table I-2. Treatment for drug-sensitive and drug-resistant tuberculosis	4
Table I-3. In vitro IC <sub>50</sub> values (μM) of rifamycins and benzoxazinorifamycins against <i>M. tuberculosis</i> WT and Rif <sup>R</sup> RNA polymerase	7
Table I-4. Bacterial RNA polymerase inhibitors and their mechanism of action	9
Table II-1. List of oligonucleotides used	14
Table II-2. Sequences of oligonucleotides for nanocircle preparation	17
Table II-3. List of plasmids used	20
Table II-4. Average fluorescence values of RiboGreen and MGA detection (RCT vs Plasmid-based transcription)	25
Table II-5. Fluorescence signal of plasmid-based transcription assay with varying DNA templates	28
Table III-1. List of oligonucleotides used	37
Table III-2. List of plasmids used in this study	40
Table III-3. In vitro activity of selected compounds from HTS against <i>E. coli</i> and MTB RNAPs	53
Table III-4. Compounds with similar activity across WT and Rif <sup>R</sup> RNAP	54
Table III-5. Compounds with reduced activity against <i>E. coli</i> H526Y RNAP (cluster 6)	55
Table IV-1. List of oligonucleotides used for FID Assay	62
Table IV-2. MICs of compounds active vs <i>M. tuberculosis</i> H <sub>37</sub> R <sub>v</sub>	67
Table IV-3. IC <sub>50</sub> values (μM) of 193157 analogs against <i>E. coli</i> WT RNAP	69
Table IV-4. Activity of cyanopyrimidines +/- nitrile in plasmid-based transcription assay	71
Table IV-5. Structure-activity relationship table of cyanopyrimidines against <i>E. coli</i> and MTB WT RNAPs	74
Table V-1. List of oligonucleotides used	95

Table V-2. In vitro IC <sub>50</sub> values (μM) for rifampin, rifalazil, and bxRifs b-d in rolling circle transcription assay	99
Table V-3. RNAP IC <sub>50</sub> values (μM) for rifampin in plasmid-based transcription assay	101
Table V-A1. Determination of endogenous <i>E. coli</i> WT RNAP in Rif <sup>R</sup> mutant purifications	115
Table V-A2. In vitro activity of RifB analogs against <i>E. coli</i> and MTB RNAPs	117
Table V-A3. In vitro activity of benzoxazinorifamycins Series A against <i>E. coli</i> and MTB RNAPs	118
Table V-A4. In vitro activity of benzoxazinorifamycins Series B against <i>E. coli</i> and MTB RNAPs	119
Table V-A5. In vitro activity of benzoxazinorifamycins Series C against <i>E. coli</i> and MTB RNAPs	119

## Abstract

### Identification and Evaluation of Bacterial RNA Polymerase Inhibitors Using a Novel Plasmid-based Transcription Assay

by

Nathan Thomas Scharf

Chair: George A. Garcia

Tuberculosis (TB) is a global health problem caused by *Mycobacterium tuberculosis* with 8-10 million new cases each year according to the World Health Organization. The current treatment includes a 6-9 month treatment with four different drugs. Due to the long treatment time and lack of adherence to the treatment regimen, there is a rise in the number of multi-drug resistant strains of *Mycobacterium tuberculosis* (MDR-TB). Rifampin, a long-time staple in TB treatment, is an effective drug against TB which acts by inhibiting the bacterial enzyme RNA polymerase (RNAP). However, there are problems of resistance to rifampin due to mutations in the gene encoding for RNAP and rifampin is a very effective inducer of CYP450s. It is of interest to develop potent inhibitors against bacterial RNAP that are effective against both the wild-type *M. tuberculosis* RNAP and common mutant RNAPs and do not induce CYP450s.

Nucleic acid aptamers are very useful oligonucleotides that bind specifically to a target molecule based on the nucleotide sequence. An *in vitro* plasmid-based RNAP transcription assay was developed and adapted to high-throughput screening, in which a malachite green aptamer

(MGA) is used as the detection method for RNA transcription. Compounds identified were evaluated for activity against a panel of 10 bacterial RNAPs (*E. coli* and MTB - WT and Rif<sup>R</sup> RNAPs), and active scaffolds were further studied resulting in identification of a cyanopyrimidine scaffold for bacterial RNAP inhibition.

Previous studies have shown benzoxazinorifamycins to have improved activity compared to rifampin in vitro against Rif<sup>R</sup> RNAP enzymes. Additionally, benzoxazinorifamycins have decreased induction of the human pregnane X receptor (hPXR), which leads to expression of CYP3A4 and drug-drug interactions for patients taking rifampin concurrently with HIV medications. Novel benzoxazinorifamycins were evaluated for in vitro activity against WT and Rif<sup>R</sup> bacterial RNAPs (*E. coli* and MTB) using the plasmid-based transcription assay. We hypothesize that identification of rifamycins that do not activate hPXR and are still active against *M. tuberculosis* RNAP and *M. tuberculosis* in culture will allow for the development of anti-TB drugs that can be taken concurrently with HIV medications, as TB-HIV coinfection is a global problem.

## Chapter I

### Introduction

Despite numerous advances over the last several decades, infectious diseases remain a global health problem. Infectious diseases are the third leading cause of death globally (~10% of global deaths), behind cardiovascular disease and cancer.<sup>1</sup> Tuberculosis has been the leading cause of death in infectious diseases since 2015, ahead of HIV/AIDS and diarrheal diseases.<sup>2</sup> Since the golden age of antibiotics (approx. 1950 – 1960), a time when approximately half of antibiotics in use today were discovered, research into and development of antibiotics significantly dropped.<sup>3</sup> This is indicated by the number of antibiotics developed in each five year period from 1980 to 2017 (Figure I-1).<sup>4,5</sup> However, over the past few decades, there has been a growing number of cases of antibiotic resistant infections.<sup>6-8</sup> There are several factors that can contribute

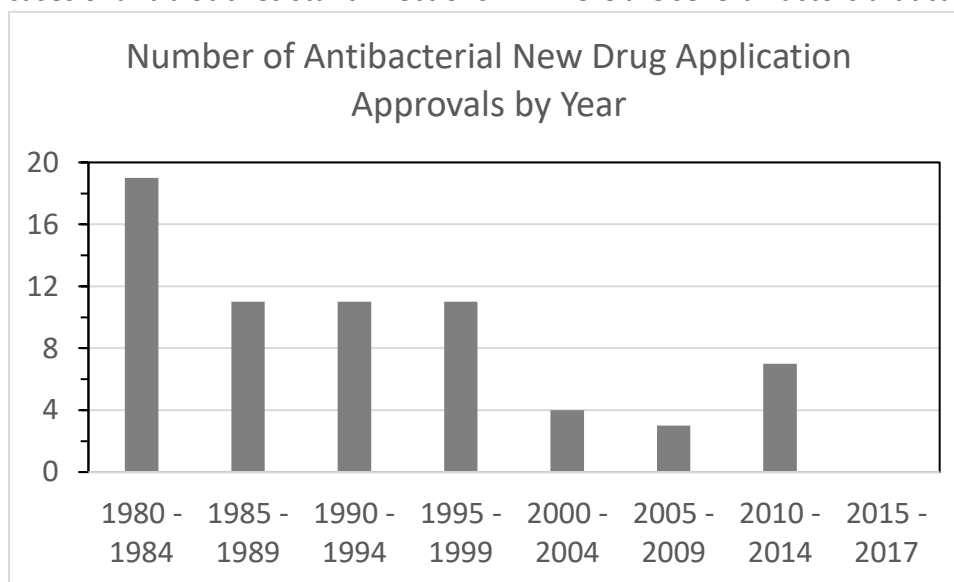


Figure I-1. Number of new antibiotics approved by year from 1980 – 2017. Figure is adapted from the CDC Antibiotic Resistance report. The drugs are limited to systemic agents (topical, ophthalmic solutions, antibodies, etc are not included). Data is from the FDA Center for Drug Evaluation and Research.

to antibiotic resistance, which include misuse of antibiotics (not completing full regimen, use in agriculture, inappropriate prescribing, etc.).<sup>4,8</sup> Due to the growing issue of antibiotic resistance coupled with the lack of novel antibiotics entering the market, there is a greater need for antibiotics that can combat the issues of resistance that are appearing in the clinic.<sup>8-10</sup>

Bacteria can develop resistance to antibiotics in a variety of ways. Major mechanisms of resistance include changes in membrane permeability or antibiotic residence time (efflux), modification/degradation of the antibiotic, or modification of bacterial target to reduce affinity.<sup>11,12</sup> It is essential to adequately understand the mechanism(s) to which bacteria become resistant to a particular antibiotic in order to overcome the resistance.

### **Tuberculosis and *Mycobacterium tuberculosis***

Tuberculosis (TB) is an infectious disease caused by the pathogen *Mycobacterium tuberculosis*, and is one of the largest global health problems, infecting an estimated 8 – 10 million new people each year (10.4 million new cases in 2015). Discovery of several anti-TB drugs initially led to an effective cure; however, tuberculosis re-emerged during the AIDS epidemic in the 1980s.<sup>13</sup> After this resurgence of TB, the World Health Organization declared TB a global health emergency in 1993.<sup>14</sup> TB can exist in two forms, a latent form or an active form. In the latent form of TB infection, *M. tuberculosis* has elicited an immunological response where the bacterial replication is under control and the affected person cannot infect others, does not feel sick, and does not have any symptoms.<sup>2,15,16</sup> A tuberculin skin test is used to detect latent TB, however it frequently yields false positive results among people vaccinated with the Bacillus Calmette-Guérin vaccine and it does not distinguish between latent and active TB.<sup>17</sup> In an active TB infection, *M. tuberculosis* bacteria overcome the immune system suppression and remain in

an actively replicating phase, manifesting in the disease tuberculosis – indicated by coughing for weeks, chest pain, weight loss, etc.<sup>16,18</sup> The likelihood of activation from latent TB to the active form is an estimated 5 to 15%, with a higher likelihood in patients coinfecting with HIV.<sup>19</sup> In 2014, there were an estimated 1.5 million deaths caused by TB (0.4 million with HIV coinfection). The global health implications of tuberculosis and *M. tuberculosis* make it an infectious disease of great interest and importance.

### Tuberculosis Treatment and Drug-Resistant Tuberculosis

The current treatment regimen for tuberculosis is a 6- to 9-month time course of four different first-line anti-tuberculosis drugs: rifampin, isoniazid, ethambutol, and pyrazinamide (Table I-1).<sup>20-</sup>  
<sup>23</sup> The global treatment success rate for newly diagnosed TB was 83% in 2014.<sup>2</sup> The duration of treatment and number of drugs required for cure of tuberculosis has been tested thoroughly.<sup>24,25</sup>  
 The recommended drug regimen by the World Health Organization is two months of the four

Table I-1. Drugs used for the treatment of tuberculosis

<b>Drug Name</b>	<b>Mode of Action</b>
<b><i>First-Line</i></b>	
Isoniazid	Inhibition of mycolic acid synthesis (cell wall)
Rifampin	Inhibition of bacterial RNA synthesis
Ethambutol	Inhibition of arabinogalactan synthesis (cell wall)
Pyrazinamide	Disruption of membrane potential
<b><i>Second-Line</i></b>	
Aminoglycosides	Inhibition of protein synthesis
Linezolid	Inhibition of protein synthesis
Ethionamide/prothionamide	Inhibition of mycolic acid synthesis (cell wall)
Delamanid	Inhibition of mycolic acid synthesis (cell wall)
Cycloserine/Terizidone	Inhibition of peptidoglycan synthesis (cell wall)
Fluoroquinolones	Inhibition of DNA synthesis
4-aminosalicylic acid	Inhibition of dihydrofolate reductase
Bedaquiline	Inhibition of ATP synthesis

first-line anti-tuberculosis drugs, followed by four months of rifampin and isoniazid, with an average cost of treatment of \$40 per person.<sup>2</sup>

Due to the emergence and prevalence of resistant strains of *M. tuberculosis*, additional therapy regimens have been designed and implemented. Strains that are resistant to rifampin and isoniazid are classified as multidrug-resistant TB (MDR-TB). The current recommended treatment for MDR-TB and rifampin-resistant (Rif<sup>R</sup>) TB is a 9- to 12-month treatment with four drugs (costing approximately \$1000 per person), which was reduced from an 18- to 24-month treatment.<sup>2</sup> The therapies for MDR-TB include some of the second-line antibiotics (Table I-1) that are less effective, administered intravenous or intramuscular, and have adverse side effects – with a 52% global treatment success rate.<sup>26–33</sup> It is estimated that almost 10% of MDR-TB strains have additional resistance to fluoroquinolones and aminoglycosides and are classified as extensively drug-resistant TB (XDR-TB), with an estimated treatment success rate of only 28% (Table I-2).<sup>2</sup> There have been some reports of tuberculosis cases that respond to few of the first and second line antibiotics, and have been referred to as totally drug-resistant TB.

Table I-2. Treatment for drug-sensitive and drug-resistant tuberculosis

Drug Susceptibility	Treatment Duration	Success Rate	Treatment Cost*	Resistance Profile
Drug-sensitive	6 - 9 months	83%	\$40	Susceptible
Multidrug-resistant	9 - 12 months	52%	\$1,000	Rifamycins, isoniazid
Extensively drug-resistant	18 - 24 months	28%	>\$5000	Rifamycins, isoniazid, aminoglycosides, fluoroquinolones
Totally drug-resistant	N/A	N/A	N/A	Most of 1st and 2nd line drugs

\*Treatment costs are the global cost of the drugs. Other associated medical costs are not included (hospitalization, etc)

Only two novel drugs have been developed and approved for tuberculosis recently, delamanid and bedaquiline, with bedaquiline being the first anti-TB drug with a novel mechanism



in four decades.<sup>34</sup> Despite the recent development and implementation of these drugs in 2014, resistant isolates have already been identified in the clinic.<sup>35–37</sup> Due to the lack of new drugs and the emerging issues of resistance, it is essential to investigate potential classes of molecules that can overcome these issues of resistance and are effective against *M. tuberculosis*. Reinvestigating previously effective drug targets is an approach that has proven to be successful in some cases for antibiotic drug discovery.<sup>38,39</sup>

### **Rifamycins, Rifampin, and Rifamycin Resistance**

The rifamycins are a class of natural product antibiotics derived from the bacteria *Amycolatopsis rifamycinica*.<sup>40</sup> Subsequent studies identified rifamycin SV (Figure I-2) and other derivatives of rifamycin B to have good activity against *M. tuberculosis* and gram-positive bacteria.<sup>21,41</sup> After introduction of rifamycin SV for treatment, a semisynthetic campaign was undertaken to obtain a compound with better oral absorption, prolonged antibacterial levels in the blood, and improved activity against *M. tuberculosis* and gram-negative bacteria.<sup>21,42</sup> This eventually led to the development and implementation of rifampin for the treatment of TB, reducing the TB treatment time from 18 to 9 months when used in combination therapy.<sup>43,44</sup> Rifampin and other rifamycins are bactericidal against *M. tuberculosis* and other pathogens by inhibiting the bacterial DNA-dependent RNA polymerase (RNAP), inhibiting initiation and preventing formation of nascent RNA longer than 2 to 3 nucleotides.<sup>45</sup>

The extensive chemical modifications made on the rifamycin molecule resulted in identifying the following relationships between rifamycin structure and its antibacterial activity.<sup>21</sup>

- Changes in the ansa chain (C-15 to C-29) generally yields less active rifamycins

- Substitution or elimination of hydroxyl groups at C-21 and C-23 results in large decreases in activity
- Hydroxy group at C-8 is essential for antimicrobial activity
- Both quinone and hydroquinone systems are active for the naphthalene core (keto and hydroxy groups at C-1/C-4 had similar activity)
- Compounds with a free carboxyl group are less active due to poor penetration
- Substitutions at C-3 and C-4 are well tolerated

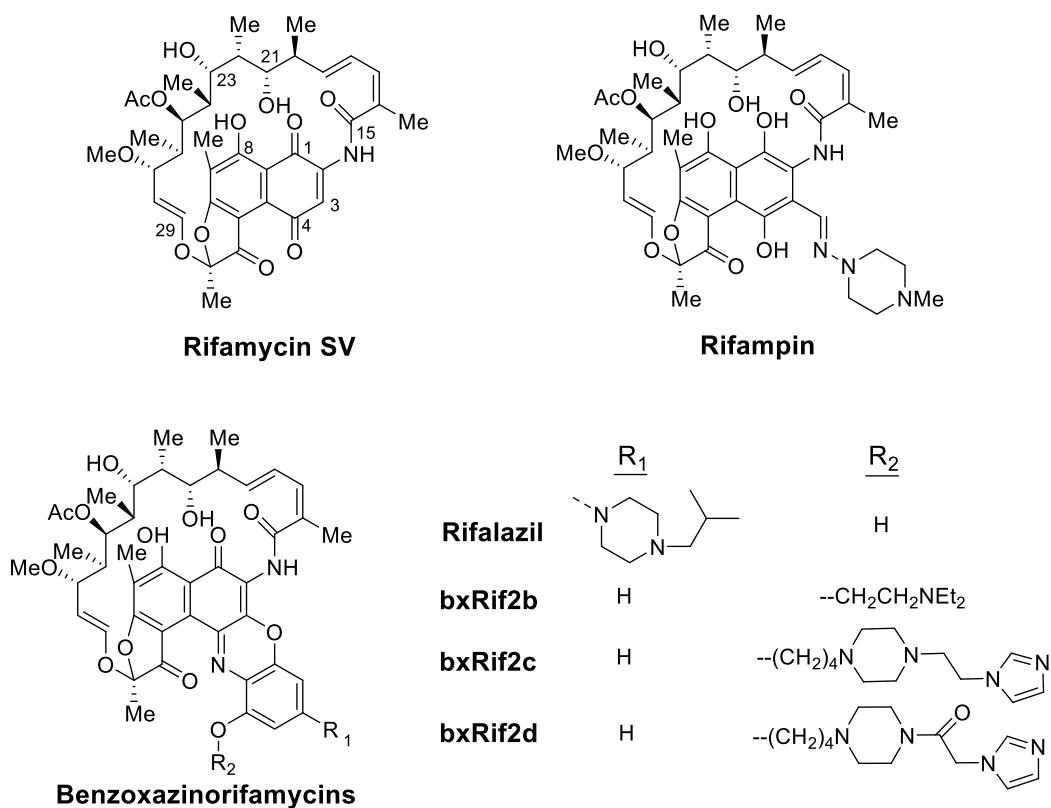


Figure I-2. Structures of rifamycins of particular interest. Numbering referenced in text is shown on rifamycin SV.

Table I-3. In vitro IC<sub>50</sub> values (μM) of rifamycins and benzoxazinorifamycins against *M. tuberculosis* WT and Rif<sup>R</sup> RNA polymerase

Compound	WT	D435V	H445Y	S450L
Rifampin	0.017	465	610	109
Rifalazil	< 0.01	541	172	117
bxRif2b	< 0.01	20	171	16
bxRif2c	< 0.01	9	437	18
bxRif2d	< 0.01	13	574	122

Despite the extensive studies to identify a rifamycin analog with favorable qualities, culminating in the identification and commercialization of rifampin, it has been shown that additional progress can still be made. Compounds have been identified with improved pharmacokinetic properties as well as partial amelioration of drug – drug interactions, which is important for the significant subpopulation of HIV patients that are coinfecting with *M. tuberculosis*.<sup>46</sup> Rifampin induces CYP3A4 via activation of the human pregnane X receptor (hPXR), which causes rifampin to exhibit drug – drug interactions with many other drugs.<sup>47</sup> The rifamycin analog rifalazil (Figure I-2) was of particular interest because it was shown to be more potent against *M. tuberculosis* than rifampin and other analogs, and it lacked toxicity in early rodent studies.<sup>48</sup> Further evaluation of rifalazil showed that it did not exhibit complete cross-resistance to rifampin and it was shown to not induce CYP3A4.<sup>49–51</sup> Rifalazil was tested in phase I and phase II clinical trials, however it was proven to be too toxic for use against TB in humans.<sup>52</sup>

Rifamycin resistance in the RNA polymerase occurs at a high rate ( $\sim 10^{-8}$ ), with an 81-bp core region of the *rpoB* gene, known as the rifampin resistance determining region (RRDR), accounting for  $\sim 96\%$  of all Rif<sup>R</sup> strains of *M. tuberculosis*.<sup>45,53–55</sup> Mutations at three residues within the RRDR (S450, H445, D435) account for approximately 84% of Rif<sup>R</sup> strains, and the most

clinically prevalent Rif<sup>R</sup> mutation (S450L) accounts for ~45% of Rif<sup>R</sup> MTB.<sup>54,55</sup> Rifalazil and other benzoxazinorifamycins (Figure I-2) have been studied more recently for in vitro activity against WT and Rif<sup>R</sup> bacterial RNAP, with improved activity against the most clinically relevant Rif<sup>R</sup> mutations compared to rifampin (Table I-3).<sup>56,57</sup> Due to the current issues of drug – drug interactions and Rif<sup>R</sup> *M. tuberculosis*, it is of interest to identify novel bacterial RNAP inhibitors that are effective against both wild-type and Rif<sup>R</sup> bacterial RNAP, do not activate hPXR, and have favorable toxicity profiles for the treatment of tuberculosis.

### **Known Bacterial RNA Polymerase Inhibitors**

Rifamycins is one of several classes of compounds that are known to inhibit bacterial RNAP. Bacterial RNAP is a multi-subunit enzyme (Core enzyme:  $\alpha\beta\beta'\omega$ ) with an associated sigma ( $\sigma$ ) factor. RNA transcription is a process that involves multiple steps for RNA to be synthesized within the cell. The sigma factor allows for the RNAP to recognize and bind the -10 and -35 consensus elements of the promoter sequence.<sup>58</sup> The transcription cycle is made up of four steps: 1) assembly of the RNAP holoenzyme from core enzyme and sigma factor, 2) formation of promoter complex and initiation of de novo RNA synthesis, 3) elongation of RNA, and 4) termination of synthesis and dissociation of RNA. Rifampin is known to inhibit the initiation of RNA synthesis; however, there are other compounds known to inhibit RNA polymerase via other mechanisms. Some of these compounds have been identified from natural product extracts, where others have been identified through high-throughput screening (Table I-4).<sup>56,59–67</sup> Including the rifamycins, the most potent compounds are natural products or natural product derivatives (GE23077, Salinamide, Sorangicin). GE23077 is the most potent inhibitor other than the rifamycins, but it is not permeable to most bacteria.<sup>63</sup>

Table I-4. Bacterial RNA polymerase inhibitors and their mechanism of action

Compound	In vitro <i>E. coli</i> RNAP IC <sub>50</sub> (nM)	Type of Inhibition
SB2 series	1000	Holoenzyme assembly
Myxopyronin	1000	Complex formation
Lipiarmycin	3000-10000	Complex formation
Rifampin	< 10	Initiation
GE23077	40	Initiation
Salinamide	200 - 500	Initiation
Sorangicin	100	Initiation
Streptolydigin	18000 (K <sub>i</sub> )	Initiation and elongation
Microcin J25	1400	Initiation and elongation
CBR703	10000	Initiation and elongation

Sorangicin is cross-resistant with the rifamycins and salinamide has the potential for off-target effects, so these compounds are not ideal for developing an anti-TB drug.<sup>62,68</sup> Some of the other RNAP inhibitors have been identified from high-throughput screens using several different assays such as a fluorescence-based assay using a non-native fluorescent nucleotide, measuring the RNAP –  $\sigma$  interaction (SB2), and using a radiolabeled nucleotide (CBR703).<sup>60,69,70</sup> The development of a fluorescence-based transcription assay using a plasmid DNA template would better recapitulate how transcription occurs within the cells, compared to the assays previously used.

### Malachite Green Aptamer

RNA aptamers allow for highly specific recognition of a molecule based on the unique three-dimensional structure of the RNA oligonucleotide. The malachite green aptamer is a 38 nucleotide RNA aptamer that binds specifically to malachite green and other triphenylmethane dyes, yielding an increase in the dye fluorescence by restricting vibrational deexcitation.<sup>71</sup> The malachite green aptamer has previously been used to detect activity of the *E. coli* RNAP core

enzyme using a transcription assay known as rolling circle transcription.<sup>72</sup> Rolling circle transcription utilizes a circular single-stranded DNA template that binds RNAP core enzyme, and does not require an associated sigma factor. The RNAP forms multiple repeats of RNA encoded by the DNA template, forming long RNAs via constitutive elongation.<sup>73</sup>

### **Research Objectives**

There is a need for novel anti-TB drugs that are effective against antibiotic resistant strains. Reinvestigation of targets for previous drugs is an attractive approach to overcome the issues of site-specific resistance. When first introduced, rifampin was a major breakthrough in the treatment for TB; however, issues of resistance have made it important for new anti-TB drugs to enter the pipeline.

The focus of this dissertation is to identify and evaluate compounds for inhibition of bacterial RNAP. Although several assays have been developed to identify and evaluate inhibitors of bacterial RNAP, the development of a novel plasmid-based transcription assay utilizing the malachite green aptamer is of interest because it would be cost-effective, a better representation of cellular transcription, and amenable to high-throughput screening. This transcription assay can also be used to evaluate novel bacterial RNAP inhibitors for their activity against *E. coli* and MTB RNAP. The questions answered in this dissertation are listed below.

- Chapter II: Can a plasmid-based transcription assay be developed for use in a high-throughput screening assay?
- Chapter III: Can novel bacterial RNA polymerase inhibitors be identified from a high-throughput screen?

- Chapter IV: Are any of the scaffolds identified from the high-throughput screen specific inhibitors of bacterial (*E. coli* and MTB) RNA polymerase that could be further developed as a potent inhibitor?
- Chapter V: Are any of our novel benzoxazinorifamycins more potent against bacterial RNAP (*E. coli* and MTB) and *M. tuberculosis* in culture than previously developed compounds?

## Chapter II

### Development and Optimization of In vitro Plasmid-Based Transcription Assay Utilizing Malachite Green Aptamer Detection

Bacterial RNA polymerase (RNAP) is the target of the first line anti-tuberculosis drug rifampin. Issues of site-specific resistance to rifampin have made it important to identify novel anti-TB drugs that are effective against resistant strains of *M. tuberculosis* (MTB) for the treatment of multidrug-resistant and extensively drug-resistant TB.

Several transcription assays have been utilized to assess RNAP activity in vitro; however, none of them are optimally suitable for high-throughput screening and follow-up. It is of interest to develop an assay that is efficient, sensitive, low cost of materials, and more accurately reflects the cellular transcription process. A novel plasmid-based transcription assay utilizing a fluorescence reporter to detect the product RNA should achieve those goals. Unlike rolling circle transcription, the plasmid-based transcription assay requires a sigma factor for recognition and binding to the DNA promoter sequence. Additionally, a plasmid-based assay utilizes a double-stranded DNA template, so unwinding of the DNA and open complex formation is required in order for initiation and elongation to occur.<sup>74</sup> A fluorescence assay is preferred over an assay utilizing a radiolabeled nucleotide because of the hazards and costs associated with radioactive materials, which would not be amenable to a high-throughput screen (HTS). A malachite green aptamer (MGA) assay was chosen over an assay that utilizes a non-native NTP analog that



releases a fluorescent molecule when incorporated into the RNA because using a native substrate is preferred and there are substantial costs associated with preparation of the nucleotide analog.

Therefore, with the goal of developing an assay that was amenable to an HTS, we incorporated the malachite green aptamer gene into a small, high-copy plasmid (pTZ18U). After a series of optimization steps, we developed a plasmid containing multiple repeats of the MGA and developed assay conditions that gave a robust signal when incubated with *E. coli* RNAP.

## Materials and Methods

### Materials

All reagents were purchased from Sigma-Aldrich (St Louis, MO), unless otherwise specified. Nucleoside triphosphates were purchased from Chem-Impex International (Wood Dale, IL). QIAprep Spin Miniprep Kit, QIAquick Gel Extraction Kit, QIAquick Nucleotide Removal Kit, and Ni-NTA agarose were from Qiagen (Valencia, CA). Yeast extract, bactotryptone, kanamycin, DTT, Corning microplates, Amicon Centrifugal Filter Units were purchased from Fisher Scientific (Hampton, NH). Carbenicillin and isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) were purchased from Gold Biotechnology (St Louis, MO). Chloramphenicol was purchased from Tokyo Chemical Industry (Tokyo, Japan). Taq DNA polymerase, T4 DNA ligase, Antarctic Phosphatase, deoxynucleoside triphosphates (dNTPs), and all restriction enzymes were purchased from New England Biolabs (Ipswich, MA). *PfuUltra* high-fidelity DNA polymerase was purchased from Agilent Technologies (Santa Clara, CA). pCR2.1-TOPO Cloning Kit, Quant-iT™ RiboGreen® RNA reagent and all synthetic oligonucleotides were purchased from Invitrogen (Carlsbad, CA), unless otherwise specified. Oligonucleotide primers used in this study are listed in Table II-1. Oligonucleotides for nanocircle preparation (Table II-2) were purchased from

Integrated DNA Technologies (Coralville, IA). Plasmids used and prepared in this study are listed in Table II-3. The Bio-Rad Protein Assay Kit was purchased from Bio-Rad (Hercules, CA). DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility. MALDI-TOF mass spectrometry was performed by the University of Michigan Chemistry Mass Spectrometry Core. GELase™ Agarose Gel-Digesting Preparation was purchased from Epicentre (Madison, WI). Seaplaque® Agarose was from Lonza (Rockland, ME).

Table II-1. List of oligonucleotides used

Oligonucleotide Name (Restriction enzyme, if any)	Sequence, 5' - 3' (Restriction site underlined, if any)	Length
prrn-Fwd (EcoRI)	<u>GAATTCT</u> CGTGGAGAACCTG	20
prrn-Rev (XhoI)	CTC <u>GAGT</u> GCCAGTCTAATACA	21
4xMGA-Fwd (XhoI)	CTC <u>GAGC</u> GGCAGCACGTCGAC	21
4xMGA-Rev (HindIII)	<u>AAGCTT</u> GCATGCCTGCAGGTCC	22
lacUV5-Fwd	TTATGCTTCCGGCTCGTATAATGTGTGGAATTGTGAGCG	39
lacUV5-Rev	CGCTCACAATTCCACACATTATACGAGCCGGAAGCATAA	39
lacCONS-Fwd	ATTAGGCACCCAGGCTTGACACTTATGCTTCCGGCTCGT	40
lacCONS-Rev	ACGAGCCGGAAGCATAAGTGTCAAGCCTGGGGTGCCTAAT	40
280Del-Fwd	CGCGGCCTTTTACGGTCCGGCACCCAGGCTTTACAC	39
280Del-Rev	GTGTAAAGCCTGGGGTGCCGGAACCGTAAAAAGGCCGCG	39
358Del-Fwd (XbaI)	CCCCGTCAAGCTCTAG <u>AAT</u> CGGGGGCTCC	29
358Del-Rev (XbaI)	GGAGCCCCGAT <u>TCTAG</u> AGCTTGACGGGG	29
MGAclone-Fwd (NheI)	<u>GCTAGC</u> GGCAGCACGTCGAC	20
MGAclone-Rev (XbaI)	GGACCGAAGTCCGCTCTAG <u>AGG</u>	22

The plasmid construct pVS10 for WT *E. coli* RNAP expression was a generous gift from Dr. Irina Artsimovitch (Ohio State University). The preparations of pVS10 *E. coli* RNAP expression plasmids for  $\beta$  subunit mutations D516V, H526Y, S531L were previously described.<sup>56</sup> The pU6+19 vector containing the four MGA repeats (pU6+19-4xMGA) was a generous gift from Dr. Marit Nilsen-Hamilton (Iowa State University).<sup>75</sup> The pFPV27-prrn plasmid was a generous gift from Dr. Jaya Tyagi (India Institute of Medical Sciences).<sup>76</sup>

## Expression and Purification of *E. coli* WT RNAP

*E. coli* WT RNAP was expressed and purified from the plasmid construct pVS10. The pVS10 plasmid encoding the four subunits for RNAP core enzyme ( $\alpha$ ,  $\beta$ ,  $\beta'$ ,  $\omega$ ) was transformed into *E. coli* BL21(DE3) and grown in 1 L 2xTY with carbenicillin at 30°C until OD<sub>600</sub> = 1.0 -1.2, at which point protein expression was induced with 1 mM IPTG. The cells were grown for an additional 4 hrs at 30°C, then harvested by centrifugation. The cell pellet was suspended in RNAP lysis buffer (50 mM Tris-HCl, 1 mM EDTA, 10 mM  $\beta$ -mercaptoethanol [ $\beta$ -ME], pH = 8.0 at 4°C) and supplemented with 2 mM PMSF. Cells were lysed by sonication, and the lysate was cleared by centrifugation (35,000 x g, 30 min, 4°C). Protein was precipitated by addition of 10% polymine P (pH = 7) to a final concentration of 0.5% polymine P, incubated at 4°C for 15 min with shaking, followed by centrifugation (6,000 x g, 20 min, 4°C). The precipitate was resuspended to homogeneity with a glass rod in 15 mL of TGEB buffer (10 mM Tris-HCl, 5% glycerol, 0.1 mM EDTA, 5 mM  $\beta$ -ME, pH = 8.0 at 4°C) containing 0.4 M NaCl and collected by centrifugation (6,000 x g, 20 min, 4°C), followed by resuspension of RNAP in 15 mL of TGEB containing 1 M NaCl to solubilize RNAP. The RNAP was then clarified by centrifugation (8,000 x g, 20 min, 4°C), and the resulting protein in the supernatant was precipitated by addition of an equal volume of saturated ammonium sulfate solution to the supernatant and incubated at 4°C with constant, gentle shaking for 1 hr. After centrifugation (10,000 x g, 20 min 4°C), the protein pellet was dissolved in TGEB containing 50 mM NaCl and dialyzed against 2 L of the same buffer overnight at 4°C. The dialyzed RNAP was incubated with 2 mL of Ni-NTA agarose with constant shaking at 4°C for 1 hr. The mixture was then applied to a gravity-flow column and washed with TGEB containing 50 mM NaCl and imidazole as follows: 5 CVs – 5 mM imidazole, 5 CVs – 10 mM imidazole, 5 CVs – 20

mM imidazole, 5 CVs – 40 mM imidazole, 5 CVs – 100 mM imidazole. Fractions containing RNAP core as identified by SDS-PAGE and Coomassie staining were pooled and concentrated using a 10 kDa MWCO Amicon Ultra-15 Centrifugal Filter Unit. Concentrated protein was applied to a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated with TGED containing 200 mM NaCl. Eluted RNAP was concentrated using a 10 kDa Amicon Ultra-15 Centrifugal Filter Unit and flash-frozen in nitrogen before storing at -80°C.

### **Expression and Purification of *E. coli* $\sigma^{70}$**

*E. coli* BL21(DE3) containing pET15b-Sig70 (WT *rpoD*) were grown in 1 L 2xTY at 30°C until OD<sub>600</sub> ~ 1.0 – 1.2, then induced with 1 mM IPTG and allowed to grow for an additional 4 hrs at 30°C. Cells were harvested by centrifugation and resuspended in 15 mL  $\sigma^{70}$  lysis buffer (20 mM Tris-HCl, 0.4 M NaCl, 1 mM EDTA, 2 mM PMSF, 5 mM  $\beta$ -ME, 5% glycerol, pH = 8.0 at 4°C) then passed through a French press three times (20,000 psi) on ice. The lysate was cleared by centrifugation (35,000 x g, 30 min, 4°C) and the supernatant was applied to a 2-mL Ni<sup>2+</sup> column equilibrated with  $\sigma^{70}$  lysis buffer, then washed with 8 CVs of the same buffer. The column was then washed with 9 CV of  $\sigma^{70}$  wash buffer ( $\sigma^{70}$  lysis buffer + 15 mM imidazole and 50 mM NaCl), followed by elution with 6 CV of  $\sigma^{70}$  elution buffer (10 mM Tris-HCl, pH 8.0 at 4°C, 50 mM NaCl, 5% glycerol, 5 mM  $\beta$ -ME, 200 mM imidazole). Next, the protein was applied to a 5-mL Q Sepharose column equilibrated with  $\sigma^{70}$  elution buffer (without imidazole), and a linear gradient of 0.05 – 1 M NaCl over 40 CVs. Fractions containing  $\sigma^{70}$  were determined by SDS-PAGE, combined, and flash-frozen in liquid N<sub>2</sub>.

Table II-2. Sequences of oligonucleotides for nanocircle preparation

Oligonucleotide Name	Sequence, 5' - 3'	Length
Linear DNA (MG-45)	p <u>GTTACCTGGCTCTCGCCAGTCGGGATCC</u> GTTTTTA <u>GGATCCATTC</u>	45
Splint DNA (MG-45)	CCAGGTAACGAATGGATCC	19
Linear DNA (MG-52)	pGGCCACA <u>GGATCCATTCGTTACCTGGCTCTCGCCAGTCGGGATCC</u> ACGT ACC	52
Splint DNA (MG-52)	CCTGTGGCCGGTACGTGG	18

\*Underlined section in green represents DNA encoding for MG aptamer

### Preparation of DNA Nanocircle

A ssDNA nanocircle (MG-45) encoding the malachite green aptamer and linker sequence 5'–TAAAAAC–3' was prepared similarly to those previously described, circularizing a 5'-phosphorylated linear oligonucleotide with T4 DNA ligase using a complementary splint DNA (Table II-2).<sup>72</sup> In short, 5'-phosphorylated linear template (10 μM) and splint DNA (15 μM) were combined in 1x T4 DNA ligase buffer, denatured at 75°C for 10 min, then either allowed to slowly reach room temperature or placed on ice to quickly cool the reaction. BSA and T4 DNA ligase were added to final concentrations of 100 μg/mL and 4000 U/mL, and ligated overnight at 16°C. Exonuclease I was added to a final concentration of 200 U/mL and incubated at 37°C for 45 min, followed by heat inactivation. The reaction was ethanol precipitated, resuspended in water, further purified via QIAquick Nucleotide Removal Kit, and quantified on a Synergy H1 Hybrid plate reader using a Take3 plate. Purity was verified on a 16% polyacrylamide gel and visualized by staining with ethidium bromide. The weight of the nanocircle was determined by MALDI-TOF mass spectrometry. A 104 nt ssDNA nanocircle (MG-52) using the same sequence as Furukawa et al. was also prepared as described above, using the appropriate oligonucleotides (Table II-2).<sup>72</sup> The ssDNA templates were analyzed by MALDI-TOF to verify their size.

## Cloning, Preparation of MGA Plasmid Constructs

The *prn* operon (*prn* Op) was amplified by PCR from pFPV27-*prn* with oligonucleotides *prn*-Fwd and *prn*-Rev, and the DNA encoding four MGA repeats was amplified by PCR from pU6+19-4xMGA (contains a 5'-TAAAAAC-3' linker between repeats) with oligonucleotides 4xMGA-Fwd and 4xMGA-Rev. Each PCR product was subcloned using the TOPO TA Cloning Kit (pEXP-5-NT/TOPO). The resulting plasmids, TOPO-*prn* and TOPO-4xMGA, were isolated via miniprep (Qiagen) and confirmed by DNA sequencing. The *prn* and 4xMGA DNA fragments were inserted into pTZ18U by first digesting TOPO-*prn* (*EcoRI* and *XhoI*), TOPO-4xMGA (*XhoI* and *HindIII*), and pTZ18U (*EcoRI* and *HindIII*). The *prn* and 4xMGA fragments, and pTZ18U vector, were gel purified from Seaplaque® agarose gel with GELase™ according to the vendor's protocol. The three purified samples were ligated with T4 DNA ligase and then transformed into TOP10 chemically competent *E. coli*. The resulting plasmid, pTZ18U-4xMGA, was isolated via miniprep and confirmed with DNA sequencing (Figure II-1).

The *lac* wild-type (*lacWT*) promoter on pTZ18U-4xMGA was mutated to *lacUV5* via two-stage site-directed mutagenesis<sup>77</sup> with the complementary oligonucleotides *lacUV5*-Fwd and

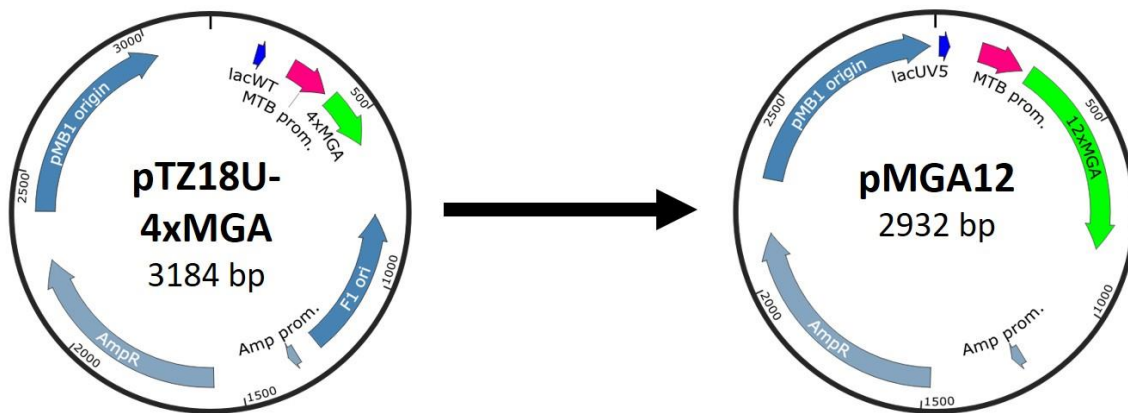


Figure II-1. Overview of modifications to prepare pMGA12 plasmid template. Original plasmid template pTZ18U-4xMGA (left) was modified by addition of eight MG repeats, mutation of promoter to *lacUV5*, removal of portion of F1 ori and DNA upstream of promoters.

lacUV5-Rev with *PfuUltra* DNA polymerase, followed by *DpnI* digestion, and transformation into TOP10 *E. coli*. The resulting plasmid, pTZ18U-4xMGA-lacUV5, was isolated via miniprep and confirmed with DNA sequencing. The *lacUV5* promoter on pTZ18U-4xMGA-lacUV5 was mutated to *lacCONS* using the same mutagenesis procedure with complementary oligonucleotides lacCONS-Fwd and lacCONS-Rev to give the plasmid pTZ18U-4xMGA-lacCONS.

A 280 bp region upstream of the *lacWT* promoter on pTZ18U-4xMGA was removed via mutagenesis using complementary oligonucleotides (280Del-Fwd and 280Del-Rev), resulting in the plasmid pTZ18U-4xMGA-280bpDel.

Additional MGA repeats were incorporated into the pTZ18U vector using an iterative cloning technique with the compatible restriction overhangs of *NheI* and *XbaI* (Figure II-2). The 4xMGA fragment was amplified by PCR from pTZ18U-4xMGA using oligonucleotides MGAclone-Fwd and MGAclone-Rev. The PCR fragment was subsequently subcloned using the TOPO Cloning Kit (pCR2.1-TOPO). The resulting plasmid, TOPO-4xMGA-*NheI*, was isolated and confirmed by DNA sequencing. pTZ18U-4xMGA was digested with *XbaI*, dephosphorylated with Antarctic

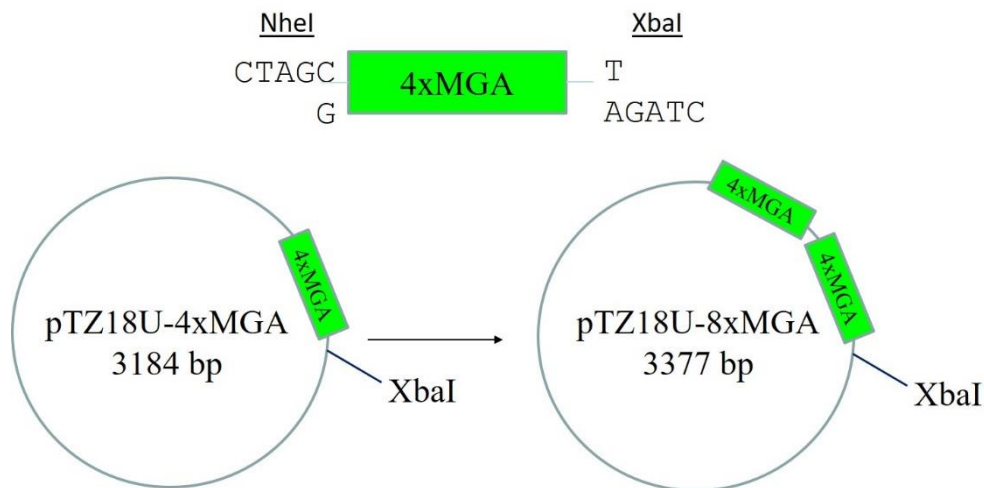


Figure II-2. Iterative cloning technique to insert additional MGA repeats. Digestion of TOPO-4xMGA-*NheI* template with *NheI* and *XbaI* results in 4xMGA insert with compatible restriction overhangs (top). Insertion of 4xMGA insert to plasmid template results in addition of MGA repeats and regeneration of *XbaI* restriction site.

Phosphatase, and then purified by gel extraction. TOPO-4xMGA-NheI was restricted with *NheI* and *XbaI*, and the 4xMGA fragment was purified by gel extraction. The 4xMGA fragment was ligated with the linearized pTZ18U-4xMGA with T4 DNA ligase, resulting in pTZ18U-8xMGA. This process was repeated to prepare a plasmid with 12 MGA repeats, pTZ18U-12xMGA.

The *lacWT* promoter on pTZ18U-12xMGA was mutated to *lacUV5* and the 280-bp region upstream of the promoters was deleted in the same manner as described above to give pMGA12+F1ori. A 358-bp region from the F1 ori was then removed from pMGA12+F1ori. First, an *XbaI* restriction site was inserted by mutagenesis using the complimentary oligonucleotides 358Del-Fwd and 358Del-Rev. The resulting plasmid was digested with *XbaI* and vector fragment was gel purified from an agarose gel using the QIAquick Gel Extraction Kit. The vector was ligated with T4 DNA ligase and then transformed into TOP10 *E. coli*. The resulting plasmid, pMGA12,

Table II-3. List of plasmids used

Plasmid Name	Description	Reference
pFPV27-prrn	pFPV27 vector containing MTB <i>prrn</i> operon	(76)
pU6+19-4xMGA	plasmid containing four repeats of MGA	(75)
TOPO-prrn	pEXP5-NT vector containing <i>prrn</i> operon	This study
TOPO-4xMGA	pEXP5-NT vector containing four MGA repeats	This study
pTZ18U-4xMGA	pTZ18U vector containing <i>prrn</i> operon and four MGA repeats	This study
pTZ18U-4xMGA-lacUV5	pTZ18U-4xMGA with <i>lacUV5</i> promoter	This study
pTZ18U-4xMGA-lacCONS	pTZ18U-4xMGA with <i>lacCONS</i> promoter	This study
pTZ18U-4xMGA-280bpDel	pTZ18U-4xMGA with 280 bp deletion upstream of <i>lacWT</i>	This study
TOPO-4xMGA-NheI	pCR2.1-TOPO vector containing four MGA repeats with <i>NheI</i> restriction site upstream	This study
pTZ18U-8xMGA	pTZ18U-4xMGA with four additional MGA repeats (8 total)	This study
pTZ18U-12xMGA	pTZ18U-8xMGA with four additional MGA repeats (12 total)	This study
pMGA12+F1ori	pTZ18U-12xMGA with <i>lacUV5</i> and 280bp deletion	This study
pMGA12	pMGA12+F1ori with 358 bp deletion downstream of MGA	This study
pVS10	Expression vector for <i>E. coli</i> RNAP core enzyme	(60)
pET15b-Sig70	Expression vector for <i>E. coli</i> $\sigma^{70}$	K. Murakami (unpublished)



was isolated and confirmed by DNA sequencing (Figure II-1). Only a portion of the F1 ori was removed because attempts to remove the entire F1 ori were unsuccessful.

### **Rolling Circle Transcription Assay**

The activity of *E. coli* RNAP was measured via rolling circle transcription in a similar fashion to that previously described.<sup>56</sup> Transcription reactions (25  $\mu$ L) were incubated at 37°C for 2 hours upon initiation by addition of NTPs under the following conditions: 40 mM Tris-HCl (pH 8.0), 50 mM KCl, 10 mM MgCl<sub>2</sub>, 0.01% Triton X-100, 8 mM DTT, 1.12 U/ $\mu$ L RNase inhibitor, 160 nM ssDNA nanocircle template, 80 nM core *E. coli* RNAP, and 500  $\mu$ M each NTP.

### **Plasmid-Based in vitro Transcription Assay**

The activity of *E. coli* RNAP was measured via plasmid-based, in vitro transcription using the plasmid templates prepared. Transcription reactions were prepared the same way as for rolling circle transcription with the following modifications: 80 nM plasmid template and 80 nM/240 nM *E. coli* RNAP/ $\sigma^{70}$  was used, unless otherwise specified in the results.

### **RiboGreen Detection of Transcription**

Transcription reactions prepared as described above were monitored by measuring the fluorescence of RiboGreen as follows. After 2 hr incubation at 37°C, 2  $\mu$ L of the reactions were quenched on ice with 73  $\mu$ L of 10 mM Tris/1 mM EDTA (pH = 7.5). Then, 75  $\mu$ L of 1:200 RiboGreen in water was added to each well, measuring fluorescence (Ex/Em = 480nm/520nm) after 5 min incubation.

### **Malachite Green Detection of Transcription**

Transcription reactions prepared as described above were monitored by measuring the fluorescence of malachite green as follows. After 2 hr incubation at 37°C, the reaction was halted

by incubating on ice and adding 75  $\mu\text{L}$  of cold 40  $\mu\text{M}$  malachite green in water. After 5 min on ice, malachite green fluorescence was detected at excitation and emission wavelengths of 628 nm and 660 nm, respectively.

### **Inhibition of *E. coli* WT RNAP by Rifampin**

The rifampin  $\text{IC}_{50}$  value on *E. coli* WT RNAP was determined using the plasmid-based transcription assay. Reactions of 50  $\mu\text{L}$  for *E. coli* WT RNAP were prepared in a 96-well plate (Corning 96-Well Half Area Black Flat Bottom Polystyrene NBS™ microplate) as follows. In each well, 24  $\mu\text{L}$  of a 2x *E. coli* reaction buffer solution [80 mM Tris-HCl (pH = 7.5 at 23°C), 100 mM KCl, 20 mM  $\text{MgCl}_2$ , 0.02% Triton X-100, 16 mM DTT, 5 nM pMGA12, 1 mM NTPs] was added, 2  $\mu\text{L}$  of DMSO or varying concentrations of two-fold serially diluted rifampin in DMSO (final, 2.5 – 80 nM), followed by 24  $\mu\text{L}$  of a 2x RNAP solution (10 nM RNAP, 30 nM  $\sigma^{70}$ ). Reactions were incubated at 37°C for 2 hrs, followed by addition of 50  $\mu\text{L}$  of 150  $\mu\text{M}$  MG in water, and incubation on ice for 10 min. Fluorescence was measured at excitation and emission wavelengths of 628 and 660 nm, respectively. Data were normalized to percent activity (DMSO = 100% activity, no RNAP = 0% activity), then fit by nonlinear regression to Eq. (1) where  $M_0$  = log of compound concentration,  $M_1$  = log of  $\text{IC}_{50}$ ,  $M_2$  = hill slope,  $M_3$  = lower limit of the curve, and  $M_4$  = upper limit of the curve.

$$(1) \quad y = M_3 + \frac{(M_4 - M_3)}{(1 + 10^{(M_0 - M_1) \times M_2})}$$

## Results

### Preparation of ssDNA Nanocircle

To evaluate the difference between RiboGreen and malachite green detection, single-stranded DNA nanocircles were prepared with either a 52 or 45 nucleotide (nt) sequence to recapitulate what has been done previously<sup>72</sup> or to be analogous to the plasmid template, respectively. Two different sized nanocircles were prepared by modifying the preparation procedure, with a larger nanocircle forming (Figure II-3, Lane 6) when allowed to cool slowly during oligonucleotide annealing.

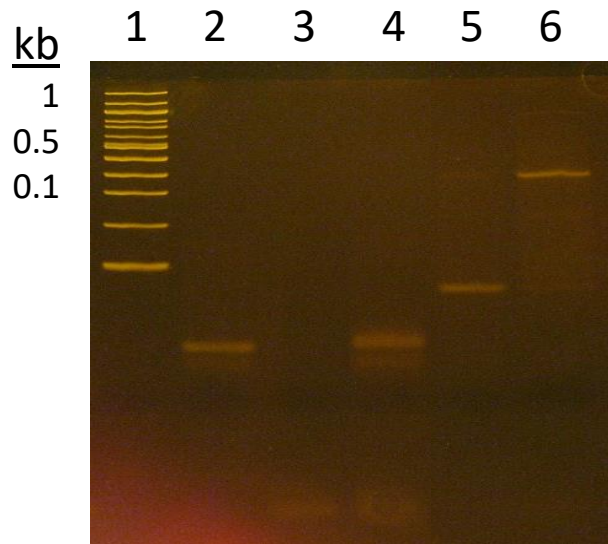


Figure II-3. 16% PAGE of synthesized DNA nanocircles. Lane 1: 100 bp ladder, Lane 2: linear DNA (MG-45), Lane 3: splint DNA (MG-45), Lane 4: hybridized linear and splint DNA (MG-52), Lane 5: quick cool reaction, Lane 6: slow cool reaction

The molecular weights of the nanocircles were determined by MALDI-TOF. It was determined that the larger nanocircle prepared contained two repeats of the 45-nt sequence, making it 90-nt in length with a molecular weight of  $\sim 27.8$  kDa (Figure II-4). The peak at 17.6 kDa is potentially explained by residual splint DNA if three pieces of splint DNA annealed together at the ends ( $MW = 5821 \times 3 = 17463$  Da) The 90-nt nanocircles were more active than the 45-nt nanocircles and were therefore used for the transcription assays. The nanocircles MG-45 and

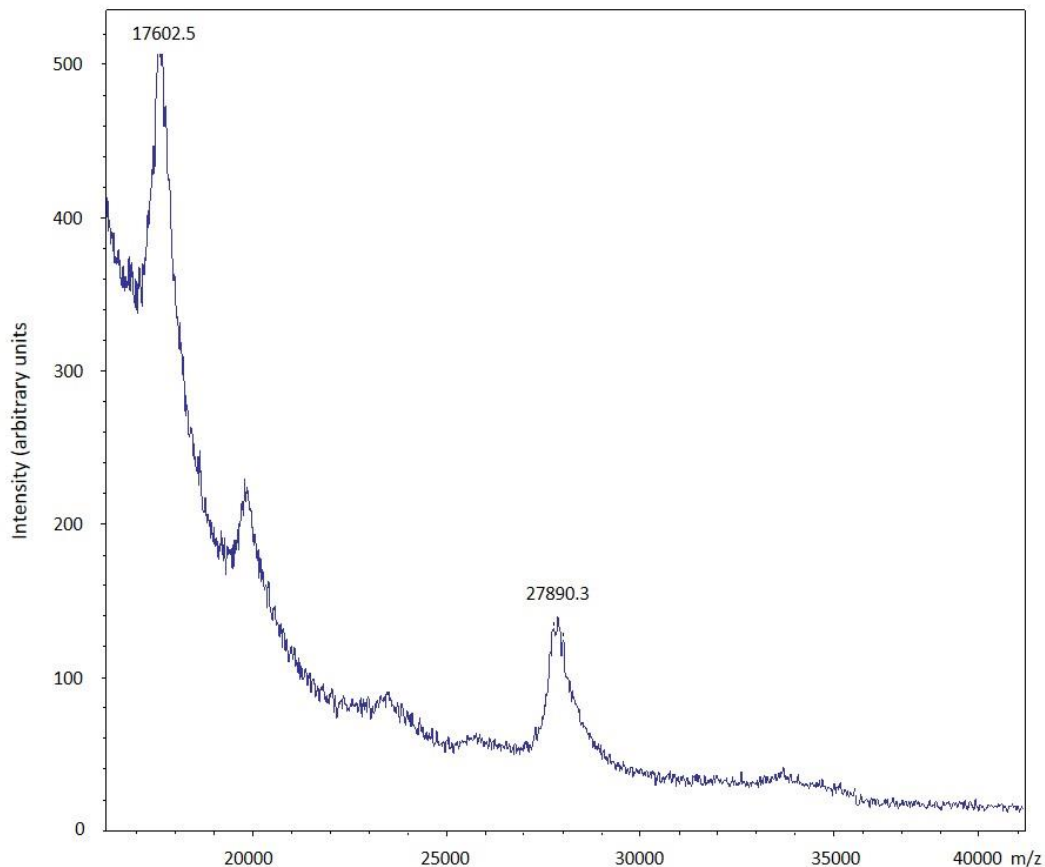


Figure II-4. Mass spectra of circularized MG-45. Mass spectra was obtained by MALDI-TOF on a Bruker Autoflex Speed mass spectrometer in positive ion linear mode, calibrated with a mixture of proteins in sinapinic acid. Sample was rinsed with 0.1 M triethylammonium acetate.

MG-52 refer to the larger (90- and 104-nt) nanocircles. The MG-45 nanocircle was used instead of the MG-52 nanocircle because it has the same sequence as in the plasmid template, and the transcription signal for each nanocircle were similar to one another (data not shown).

### **RiboGreen vs Malachite Green Detection**

The activity of *E. coli* WT RNAP was assessed in rolling circle transcription using MG-45 and in the plasmid-based transcription assay utilizing pTZ18U-4xMGA as the plasmid template. After the transcription reactions and fluorescence measurements (RiboGreen and malachite green) of both the rolling circle transcription assay and the plasmid-based transcription assay, it was determined that the total fluorescence signal from the MGA detection was lower (Figure II-

5). However, the signal to noise (S/N) ratio for MGA compared to RiboGreen was significantly better (Table II-4). The plasmid-based assay had S/N of 1.6:1 and 11:1 for RiboGreen and malachite green, respectively. The rolling circle transcription assay had S/N of 6:1 and 460:1 for RiboGreen and malachite green, respectively. Despite the improved signal to noise for the MGA assays, the fluorescence signal was significantly lower. The MGA assays required the sensitivity on the Synergy H1 Hybrid plate reader to be higher.

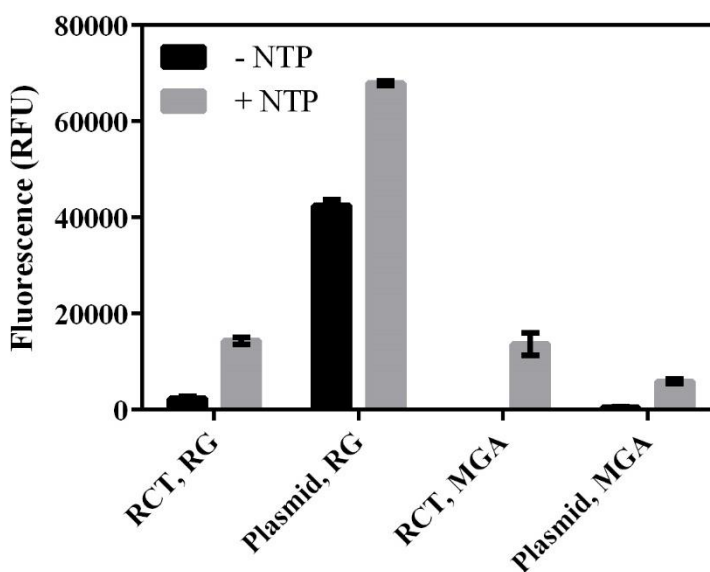


Figure II-5. Fluorescence detection of rolling circle transcription and plasmid-based transcription. Fluorescence is measured using RiboGreen (RG) or the malachite green aptamer (MGA). Fluorescence is at different instrument sensitivity values (RiboGreen = 100, MGA = 150).

Table II-4. Average fluorescence values of RiboGreen and MGA detection (RCT vs Plasmid-based transcription)

NTP	RiboGreen		MGA	
	RCT	Plasmid	RCT	Plasmid
-	2400	42500	30	540
+	14400	67900	13700	5900
Signal to Noise	6.0	1.6	460	11

## Modifications to Plasmid Template

Several modifications were made to the initial plasmid construct, pTZ18U-4xMGA, with the goal to increase the overall signal of the assay, and decrease the amount of plasmid required for a high-throughput screen. Each of the modifications made were done with at least one of these goals in mind. The size of the plasmid was decreased by the initial 280-bp deletion upstream of the promoter sequences and by the 358-bp deletion downstream of the MGA repeats. The *lacWT* promoter was mutated to the *lac* consensus promoter (*lacCONS*) as well as the *lacUV5* promoter, which has been shown to be a strong promoter within the cell.<sup>78</sup> Additional sets of four repeats were successfully cloned into the plasmid pTZ18U-4xMGA using an iterative cloning technique with the compatible restriction sites *NheI* and *XbaI*. The use of these two restriction sites allowed for the regeneration of an *XbaI* restriction site downstream of the inserted 4xMGA fragment (Figure II-2). Addition of repeats beyond 12 MGA repeats was unsuccessful. Transformations of plasmids containing numerous repeats (i.e., 12 or more) often led to the loss or removal of some repeats during the transformation and cellular replication process. The *lacUV5* and 280-bp deletion were successfully mutated onto the plasmid with 12 MGA repeats (pTZ18U-12xMGA) to give pMGA12+F1ori. After deletion of a 358-bp portion of the F1 ori downstream of the MGA repeats, the plasmid pMGA12 was prepared. pMGA12 contains the following modifications from the original plasmid: it has 8 additional MGA repeats (12 total), contains the *lacUV5* promoter sequence, and is 8% smaller (252 bp) than pTZ18U-4xMGA (Figure II-1).

## Effect of Plasmid Modifications on *E. coli* RNAP Transcription

The effects of the DNA deletions, promoter type, and number of MGA repeats were determined using *E. coli* WT RNAP in the plasmid-based in vitro transcription assay with malachite green detection. The results are shown in Figure II-6; comparing the effect of different promoters and a deletion upstream of the promoters (top left), addition of multiple sets of MGA repeats (top right), and deletion of DNA downstream of the MGA repeats (bottom). The *lacUV5* promoter was shown to be approximately 1.6-fold more active in vitro than *lacWT*, while the consensus sequence (*lacCONS*) had similar activity to *lacWT* (Figure II-6, top left and Table II-5). The plasmid containing a 280-bp deletion upstream of the *lacWT* promoter (pTZ18U-4xMGA-280bpDel) was slightly more active than pTZ18U-4xMGA (3-fold, Table II-5), and was approximately 10% smaller

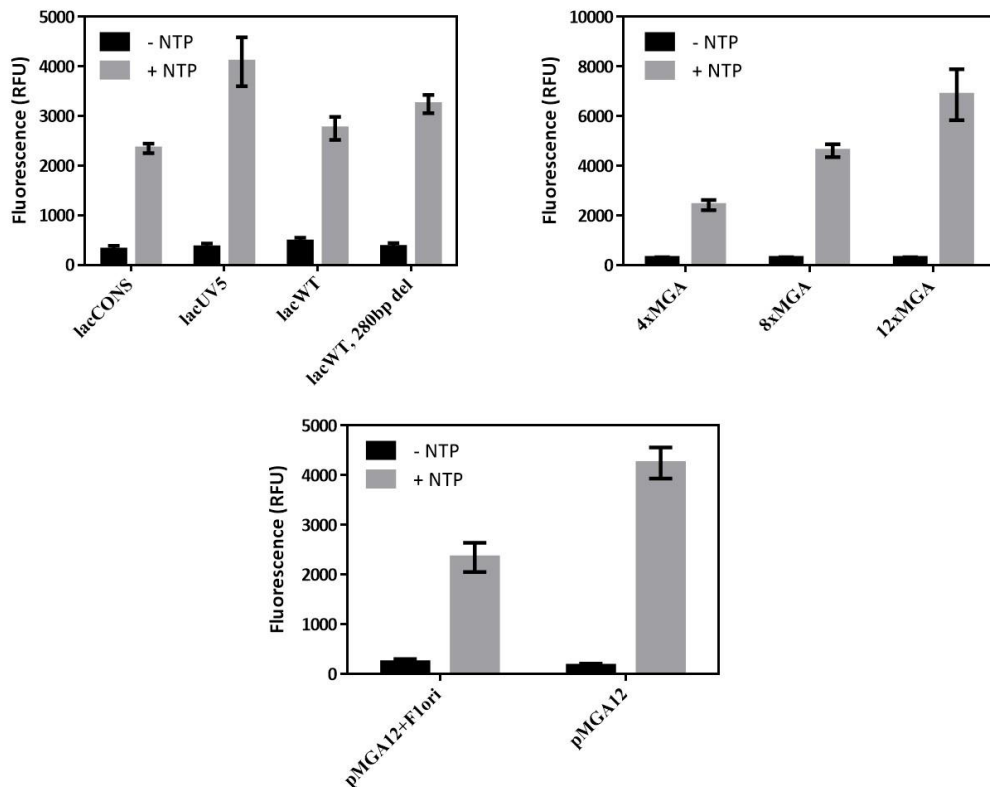


Figure II-6. Plasmid-based transcription assay using different plasmid DNA templates. Top left: Effect of *lac* promoter and 280 bp deletion upstream of promoter on *E. coli* WT RNAP transcription activity. Top right: Effect of additional MGA repeats on *E. coli* WT RNAP transcription activity. Bottom: Effect of F1 ori deletion (358 bp) downstream of MGA repeats on *E. coli* WT RNAP transcription activity.

in size. The addition of more MGA repeats onto the plasmid template via the iterative cloning technique resulted in a linear increase in the fluorescence signal from the transcription assay, using *E. coli* WT RNAP (Figure II-6, top right). The plasmid templates with 8 and 12 total MGA repeats (pTZ18U-8xMGA and pTZ18U-12xMGA), resulted in approximately a 1.8- and 2.8-fold increase relative to pTZ18U-4xMGA, respectively (Table II-5). The pMGA12+F1ori and pMGA12 plasmid templates were used for transcription by *E. coli* WT RNAP, using 40 nM plasmid template and 40 nM/120 nM *E. coli* RNAP/ $\sigma^{70}$ . A lower concentration of RNAP and plasmid were used because of the improved signal to noise from the modifications made to the DNA template. The deletion of the F1 ori downstream of the MGA repeats resulted in an approximately 1.8-fold increase in activity (Figure II-6, bottom and Table II-5), and resulted in an 11% decrease in plasmid size (358 bp).

Table II-5. Fluorescence signal of plasmid-based transcription assay with varying DNA templates

DNA Template	Fluorescence (RFU)	Signal Relative to pTZ18U-4xMGA
pTZ18U-4xMGA	2500	1
pTZ18U-4xMGA-lacCONS	2300	0.9
pTZ18U-4xMGA-lacUV5	4100	1.6
pTZ18U-4xMGA-280bpDel	3200	1.3
pTZ18U-8Xmga	4600	1.8
pTZ18U-12xMGA	6900	2.8
DNA Template	Fluorescence (RFU)	Signal Relative to pMGA12+F1ori
pMGA12+F1ori	2400	1
pMGA12	4200	1.8



## Inhibition of *E. coli* WT RNAP by Rifampin

The inhibitory effect of rifampin on *E. coli* WT RNAP was evaluated using the plasmid-based transcription assay with the DNA template pMGA12. Rifampin was tested in triplicate at the concentrations specified in the Materials and Methods section. The data were plotted (log rifampin concentration (M) vs. percent activity) and fit by nonlinear regression using Graph Pad Prism 6. The IC<sub>50</sub> value determined for rifampin was 5 nM (Figure II-7). However, the amount of RNAP used in each reaction was 5 nM (the lowest concentration that could confidently be used). Under these conditions, it is possible that a 5 nM IC<sub>50</sub> value may be at the limit of detection for the IC<sub>50</sub> value, and the actual value may be less than 5 nM.

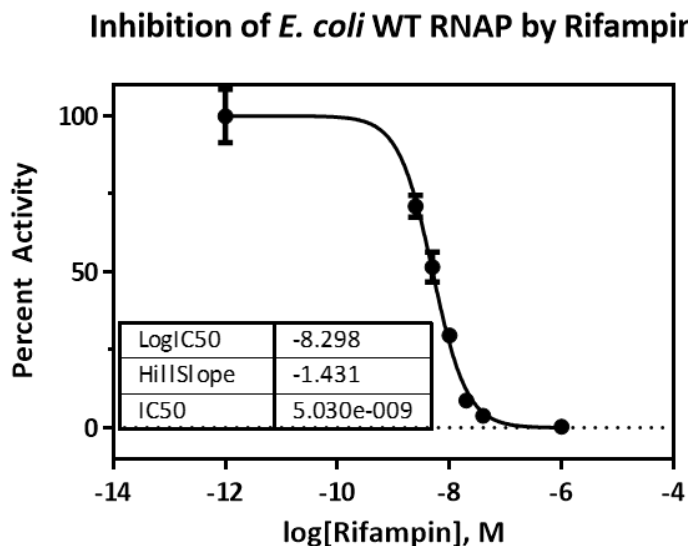


Figure II-7. Inhibition of *E. coli* WT RNAP by rifampin in plasmid-based transcription assay. Assay was tested at 5 nM RNAP and 2.5 nM pMGA12 (n=3). Data were fit by nonlinear regression using GraphPad Prism 6.

## Discussion

Bacterial RNAP is a well-validated drug target, and is an important target in the treatment of tuberculosis.<sup>21,46,58</sup> Issues of site-specific mutations to the *M. tuberculosis* RNAP have resulted in the emergence of drug-resistant cases of tuberculosis that are not susceptible to rifampin. An assay amenable to high-throughput screening using the commercially available RiboGreen has been reported; however, in addition to the cost associated with the RiboGreen dye, it requires removal and purification of RNA from the DNA template since RiboGreen fluorescence is enhanced by binding to DNA as well as RNA.<sup>79</sup> Previous assays to identify novel inhibitors of bacterial RNAP have used non-native substrates, which can be cost-prohibitive for high-throughput screening.<sup>70,80</sup>

A novel plasmid-based transcription assay has been developed that monitors the increase in fluorescence of malachite green upon binding to a 38-nucleotide RNA aptamer specific for triphenylmethane dyes.<sup>71</sup> The MGA has previously been applied to detection of *E. coli* WT RNAP activity via rolling circle transcription of a small, circular, single-stranded DNA template.<sup>72</sup> We adapted this approach to use the aptamer for detection of RNA transcribed from a double-stranded DNA template. After difficulties in detecting transcription by the MGA while using the initial plasmid template prepared (pTZ18U-4xMGA), the rolling circle transcription assay was utilized to recapitulate the assay used previously by Furukawa et al. It was determined that the sensitivity of the MGA repeats was quite low compared to RiboGreen, and a higher sensitivity was required for the Synergy H1 Hybrid plate reader (Sensitivity = 150). Based on the activity of *E. coli* WT RNAP with the original plasmid template (pTZ18U-4xMGA), the amount of plasmid and RNAP required for an HTS would be over 1 g and 100 mg, respectively. Both of these numbers

were prohibitive for an HTS, therefore it was necessary to optimize the assay in order to increase the signal and allow for less RNAP and plasmid to be used per reaction.

The plasmid pTZ18U-4xMGA was modified to decrease the size of the plasmid, increase the number of MGA repeats, and mutate the *lacWT* promoter to the *lacUV5* promoter. Decreasing the size of the plasmid was advantageous because then a smaller amount (mass) would be required for an HTS. The actual mass of plasmid required, and not number of plasmid molecules, is relevant because the capacities of DNA purification columns are based on the mass. Additionally, removal of the F1 ori resulted in a 1.8-fold increase in signal (Table II-5 and Figure II-6, bottom). Mutation to the *lacUV5* promoter instead of the *lacWT* promoter resulted in a plasmid that had higher activity in the transcription assay, as expected based on previous work studying promoter strength.<sup>78</sup>

The iterative cloning technique successfully allowed for the insertion of multiple sets of MGA repeats, while regenerating the *XbaI* restriction site. It was observed that the higher number of repeats were often instable in transformations, and it was not uncommon for plasmids with a large number of MGA repeats to lose some of these repeats, making it necessary to continually verify the plasmid sequence by DNA sequencing when making modifications or purifying more plasmid DNA. The cells used for transformation, *E. coli* TOP10, have a mutation in *recA*, which is known to delete multiple repeats of DNA.<sup>81,82</sup> However, there are RecA-independent recombination events that affect repetitive DNA sequences.<sup>83,84</sup> Therefore, it is likely that during the transformation and cell growth, additional MGA repeats can be lost through recombination mechanisms within the cell.

This assay was verified for detection of *E. coli* WT RNAP inhibition using the potent RNAP inhibitor rifampin. The assay clearly monitored concentration-dependent inhibition and recapitulated (within a factor of 2) literature values for the inhibition of *E. coli* WT RNAP.<sup>56</sup> It should be noted that the literature values for RNAP inhibition were obtained using an artificial DNA promoter in the ssDNA nanocircle, whereas the transcription assay used in this study utilized a double-stranded DNA template that requires promoter recognition and binding by an associated sigma factor.

### **Conclusions**

We have developed a robust plasmid-based transcription assay so that we can perform in vitro transcription assays with plasmid DNA purified from bacterial *E. coli* cells instead of a synthetic DNA template as is used in rolling circle transcription. This assay requires a promoter and sigma factor for DNA binding and transcription initiation, whereas the rolling circle transcription assay is single-stranded and does not require a sigma factor or promoter recognition. We are able to use the common lab reagent malachite green for RNA detection instead of the more expensive, proprietary RNA detection reagent RiboGreen. We have optimized the plasmid DNA template so that it gives a strong signal and is amenable to high-throughput screening.

## Notes to Chapter II

### Acknowledgements

I thank Dr. Irina Artsimovitch (Ohio State University) for providing the pVS10 plasmid, Dr. Jaya Tyagi (India Institute of Medical Sciences) for providing the pFPV27-prrn plasmid, and Dr. Marit Nilsen-Hamilton (Iowa State University) for providing pU6+19-4xMGA.

I would like to thank Kristen Weise, PharmD, for preparation of the initial plasmid construct pTZ18U-4xMGA., and Arrin Kontos for his work on the removal of the F1 ori from PMGA12+F1ori.

The work described in this chapter was published in *SLAS Discovery* in 2017 (Volume 22, Issue 3; pages 287 – 297).

### Abbreviations

$\beta$ -ME,  $\beta$ -mercaptoethanol; BSA, bovine serum albumin; CV, column volume; DNA, deoxyribonucleic acid; dNTP, deoxynucleotide triphosphate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HTS, high-throughput screen; IC<sub>50</sub>, half-maximal inhibitory concentration; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; MALDI-TOF, matrix assisted laser desorption/ionization-time of flight; MG, malachite green; MGA, malachite green aptamer; MTB, *Mycobacterium tuberculosis*; MWCO, molecular weight cutoff; nt, nucleotide; NTP, nucleoside triphosphate; OD<sub>600</sub>, optical density at 600 nm; ori, origin of replication; PAINS, pan-assay interference compounds; PMSF, phenylmethylsulfonyl fluoride; Rif<sup>R</sup>, rifampin-resistant; RNA, ribonucleic acid; RNAP, RNA polymerase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; S/N, signal to noise; ssDNA, single-stranded DNA; TB, tuberculosis; WT, wild-type

## Chapter III

### High-throughput Screen of *Escherichia coli* $\beta$ S531L/ $\beta$ 'V408G RNA Polymerase against 150,000

#### Compound Library

Inhibition of bacterial RNA polymerase (RNAP) has been a cornerstone of tuberculosis (TB) treatment since the introduction of rifamycins in the 1960s. Rifampin and other rifamycins are potent inhibitors of bacterial DNA-dependent RNAP. Since its introduction in the 1960s, rifampin has been effectively used to treat TB and reduced the treatment from 18 to 9 months (now a 6-month treatment).<sup>21</sup> However, rifampin and other rifamycins have several shortfalls including issues of resistance, flu-like syndrome at high doses, hepatotoxicity, and induction of cytochromes P450.<sup>46,85-87</sup> It has been suggested that if higher doses of rifampin were tolerated, the treatment duration for tuberculosis could be shortened.<sup>88</sup> Rifampin has issues of drug-drug interactions due to its induction of cytochromes P450 through the human pregnane X receptor, which is an issue due to the large TB population that is coinfecting with HIV. The issue of drug-resistant cases of TB (MDR and XDR-TB), make it important to identify novel drugs for TB that are effective in treating drug-resistant strains of *Mycobacterium tuberculosis* (MTB).

Since there are several important shortfalls to rifampin and other rifamycins, it is necessary to identify novel scaffolds that can inhibit both wild-type (WT) and rifampin-resistant (Rif<sup>R</sup>) strains of MTB RNAP. Once a scaffold or set of scaffolds have been identified to inhibit MTB RNAP, they can be optimized for activity and assessed for their potential as an antibacterial agent.

Even though there are several other known compounds that inhibit bacterial RNAP, these known compounds are not ideal for a variety of reasons as discussed in Chapter I (potency, permeability, etc.). Previous screens to identify bacterial RNAP inhibitors have used radioactive nucleotides, non-native nucleotide analogs, or measured RNAP – sigma factor interactions.<sup>60,69,70,80</sup> These assays are not ideal because they are either impractical (radioactive nucleotides), costly (non-native nucleotides), or detect a protein-protein interaction instead of transcription activity.

Studies have shown that the most clinically prevalent Rif<sup>R</sup> mutation ( $\beta$ S450L, MTB numbering), has a fitness defect.<sup>89</sup> In many clinical strains, this fitness defect is compensated for by specific secondary mutations in *rpoC*, where the most prevalent compensating mutation is V483G (MTB numbering) of the  $\beta'$  subunit.<sup>90</sup> In order to reflect what is likely the most prevalent Rif<sup>R</sup> RNAP in drug-resistant strains of TB (MDR- and XDR-TB), the *E. coli*  $\beta$ S531L/ $\beta'$ V408G RNAP (corresponding to  $\beta$ S450L/ $\beta'$ V483G in MTB numbering) was used for the high-throughput screen.

The plasmid-based transcription assay developed and optimized (Chapter II) was adapted and implemented for the high-throughput screen (HTS). The Rif<sup>R</sup> *E. coli* RNAP ( $\beta$ S531L/ $\beta'$ V408G) was screened against a 150,000-compound library to identify compounds that inhibit Rif<sup>R</sup> bacterial RNAP. The compounds were screened against the *E. coli* homolog of the clinically prevalent MTB  $\beta$ S450L/ $\beta'$ V408G RNAP because they have a large sequence identity (~50%) and we were able to purify large quantities of highly active *E. coli* RNAP, while the MTB RNAP purification was of very low yield and not as active.

## Materials and Methods

### Materials

All of the materials were the same as those specified in Chapter II. The Qiagen Giga Kit was from Qiagen (Valencia, CA). Oligonucleotide primers used in this study are listed in Table III-1. Plasmids used and prepared in this study are listed in Table III-2. ProxiPlate-384 Plus F microplates were from PerkinElmer (Waltham, MA). Fresh samples of selected compounds from the high-throughput screen were purchased from the vendors listed in Table III-3. The pMTBRP plasmids and pVS10 *E. coli* RNAP expression plasmids for  $\beta$  subunit mutations D516V, H526Y, S531L were prepared by Dr. Suman Gill as previously described.<sup>56</sup>

A 150,554-compound library was screened at the Center for Chemical Genomics at the University of Michigan. The library consisted of compounds from the Chemical Diversity set (ChemDiv-120,000 compounds), Maybridge HitFinder, MicroSource Spectrum, and NCC BioFocus libraries.

### Cloning, Preparation of *E. coli* $\beta$ S531L/ $\beta$ 'V408G RNAP Plasmid

An expression plasmid for *E. coli* S531L RNAP with the compensatory mutation  $\beta$ 'V408G was prepared from the expression plasmid pVS10(S531L) as follows by Dr. Irosha Nawarathne. The wild-type *E. coli rpoC* gene was amplified from pVS10 using oligonucleotides ECrpoC-Fwd and ECrpoC-Rev, and subcloned using the TOPO Cloning Kit (pCR2.1-TOPO). The compensatory mutation V408G (GTT – GGT) was introduced to *rpoC* via two-stage site-directed mutagenesis with complementary oligonucleotides ECrpoCMut-Fwd and ECrpoCMut-Rev.<sup>77</sup> The mutant *rpoC* gene and pVS10 (S531L) were restricted with *SbfI* and *SacII*, and ligated with T4 DNA ligase to prepare pVS10 (S531L/V408G).



Table III-1. List of oligonucleotides used

Oligonucleotide Name (Restriction enzyme)	Sequence, 5' - 3' (Restriction site underlined, if any)	Length
ECrpoC-Fwd	GGTGGTCAGCGTTTCGGGGAGATGGAAGTGTGG	33
ECrpoC-Rev	GGGATCAGACGACCCACGATAACGTTCTCTTTCAGG	36
ECrpoCMut-Fwd	GCGAAGAAGCTGTCGGTTGGGATATCCTGGACGAAGTTATCCG	43
ECrpoCMut-Rev	CGGATAACTTCGTCAGGATATCCCAACCGACAGCTTCTTCGC	43
MTBrpoA-Fwd (BamHI)	GAGAGGATCCTATGCTGATCTCACAGCGCCCCACCCTGTCC	41
MTBrpoA-Rev (HindIII)	GGGGAAGCTTGGCTAAAGCTGTTTCGGTTTCGGCGTAGTCCTGCTCGTC G	49
MTBrpoB-Fwd (NcoI)	GAACCCCATGGCAGATTCCCGCCAGAGCAAAACAGCCGC	39
MTBrpoB-Rev (BamHI)	GATCGGATCCTTAATGGTGATGGTGATGGTGATGGTGATGGTGCGCA AGATCCTCGACAC	60
MTBrpoC-Fwd (NdeI)	GGGAGTCATATGCTCGACGTCAACTTCTTCGATGAACTCC	40
MTBrpoC-Rev (EcoRV)	CGTTATTTGAGGATATCCTAGCGGTAGTCGCTGTAGCCG	39
MTBrpoZ-Fwd (NdeI)	ATTACATATGAGTATCTCGCAGTCCGACGCG	31
MTBrpoZ-Rev (KpnI)	GAGTGGTACCCTACTCGCCCTCGGTGTGC	30
MTBsigA-Fwd (NcoI)	GAAGGGGTGTCCATGGCAGCGACCAAAGCAAGCACGGCGACCGATG AGCC	50
MTBsigA-Rev (MfeI)	GCCTCGGCGGCAATTGTCAGTCCAGGTAGTCGCGCAGGACCTGTGAG CGG	50
MTBrpoCMut-Fwd	CCCCAAGGCTGGGATGTGCTCG	23
MTBrpoCMut-Rev	CGAGCACATCCCAGCCTTGGGGG	23

### Cloning, Preparation of MTB RNAP Plasmids

MTB RNAP expression vectors pACYCDuet-rpoA-rpoZ, pETDuet-rpoB-rpoC, and pRSF-SigA were prepared by Maxwell Stefan. The pACYCDuet-rpoA-rpoZ and pETDuet-rpoB-rpoC vectors were prepared using the same approach as Banerjee et al.<sup>91</sup>, amplifying and inserting the MTB genes using appropriate primers and restriction enzymes (Table III-1) from a construct previously prepared in our lab (pMTBRP-5).<sup>56</sup> The *sigA* gene encoding MTB SigA was amplified from genomic DNA of *M. tuberculosis* H<sub>37</sub>R<sub>v</sub> using oligonucleotides MTBsigA-Fwd and MTBsigA-Rev. pRSF-SigA was then prepared by restricting the *sigA* PCR fragment and pRSFDuet-1 with *NcoI* and *MfeI*, followed by phosphatase of pRSF vector, and ligation of the *sigA* gene and pRSF vector.

Expression vectors for mutant RNAPs (D435V, H445Y, S450L) were prepared by amplifying the mutant *rpoB* genes from previously prepared constructs (pMTBRP-6 thru 8) and inserting into pETDuet-rpoB-rpoC using the restriction enzymes *NcoI* and *BamHI*.<sup>56</sup> The expression plasmid for MTB  $\beta$ S450L/ $\beta$ 'V483G RNAP was prepared by mutating the *rpoC* gene subcloned in a TOPO vector to encode V483G (GTG – GGC) via two-stage site-directed mutagenesis using complementary oligonucleotides MTB $\beta$ rpoCMut-Fwd and MTB $\beta$ rpoCMut-Rev, then inserting into pETDuet-rpoB(S450L)-rpoC. pMCSG7-SigA was prepared previously.<sup>57</sup>

### **Purification of *E. coli* Rif<sup>R</sup> RNAPs**

The *E. coli* D516V, H526Y, and S531L RNAPs were purified as described for *E. coli* WT RNAP in Chapter II, using the appropriate pVS10 expression vectors for each *rpoB* mutant. *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP used for the HTS was purified by Dr. Vadim Molodtsov (Penn State University) as described for *E. coli* WT RNAP with the following modifications. Cultures of 3 L *E. coli* BL21(DE3) containing pVS10(S531L/V408G) were grown, induced, and harvested as described in Chapter II. After precipitation by polymine P, the precipitate was resuspended to homogeneity with a glass rod in 30 mL of TGEB buffer (10 mM Tris-HCl, 5% glycerol, 0.1 mM EDTA, 5 mM  $\beta$ -ME, pH = 8.0 at 4°C) containing 1 M NaCl to solubilize RNAP, then centrifuged (8,000 x g, 20 min, 4°C) and further purified by ammonium sulfate precipitation and dialysis as described in Chapter II.

The dialyzed RNAP was applied to a 50-mL cation-exchange Bio-Rex 70 column. The column was washed with 2 column volumes (CVs) of TGEB containing 100 mM NaCl, and protein was eluted with 0.1 – 1 M NaCl gradient over 5 CVs. The eluate was applied to a 10-mL Ni-NTA column equilibrated with TGEB containing 1 M NaCl at a flow rate of 0.5 mL/min. The column

was washed with 4 CVs of same buffer, then 2 CVs of TGEB containing 50 mM NaCl and 10 mM imidazole. RNAP was eluted with TGEB containing 50 mM NaCl and 100 mM imidazole. The sample was applied to a 20-mL Q Sepharose column equilibrated with TGED (TGEB with 5 mM DTT instead of  $\beta$ -ME) containing 50 mM NaCl at a flow rate of 1 mL/min. The column was washed with 2 CVs TGED, and RNAP was eluted with 0.1 – 1 M NaCl gradient over 10 CVs. Eluate was concentrated at 4°C in a Vivaspin® 20 Centrifugal Concentrator to ~ 3 mL and applied to a HiLoad 16/600 Superdex 200 equilibrated with RNAP storage buffer (10 mM Tris-HCl, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, pH = 8.0 at 4°C). The protocol above was repeated until a sufficient amount of enzyme was obtained for the HTS (85 mg). Eluate was collected and concentrated the same as above, then flash-frozen in nitrogen and stored at -80°C.

#### **Purification of *E. coli* $\sigma^{70}$**

The *E. coli*  $\sigma^{70}$  was purified by Dr. Vadim Molodtsov as follows. *E. coli* BL21(DE3) containing pET15b-Sig70 (WT *E. coli* *rpoD*) were grown in 3 L 2xTY at 30°C until OD<sub>600</sub> ~ 0.5, then induced with 1 mM IPTG and allowed to grow for an additional 2 hrs at 22°C. Cells were harvested by centrifugation and resuspended in 30 mL  $\sigma^{70}$  lysis buffer (20 mM Tris-HCl, 0.4 M NaCl, 1 mM EDTA, 2 mM PMSF, 5 mM  $\beta$ -ME, 5% glycerol, pH = 8.0 at 4°C) then sonicated on ice. The lysate was cleared by centrifugation (35,000 x g, 30 min, 4°C) and the supernatant (30 mL) was applied to a 10-mL Ni<sup>2+</sup> column equilibrated with  $\sigma^{70}$  lysis buffer + 20 mM imidazole, then washed with 20 CVs of the same buffer. Protein fractions (25 mL) were eluted with 50 mL  $\sigma^{70}$  elution buffer (10 mM Tris-HCl, pH 8.0 at 4°C, 200 mM NaCl, 5% glycerol, 0.1 mM EDTA, 5 mM  $\beta$ -ME, 200 mM imidazole). Next, the protein was diluted to reduce the concentration of NaCl to 100 mM and applied to a 5-mL Q Sepharose column equilibrated with  $\sigma^{70}$  elution buffer (without imidazole),

and a linear gradient of 0.1 – 0.6 M NaCl over 20 CVs. Fractions containing  $\sigma^{70}$  were determined by SDS-PAGE, combined (60 mg in 20 mL) and flash-frozen in liquid N<sub>2</sub>.

Table III-2. List of plasmids used in this study

Plasmid Name	Description	Reference
pMGA12	pMGA12+F1ori with 358 bp deletion downstream of MGA repeats	This study
pVS10	Expression vectors for <i>E. coli</i> RNAP core enzymes	(56)
pET15b-Sig70	Expression vector for <i>E. coli</i> Sig70	K. Murakami (unpublished)
pMCSG7-SigA	Expression vector for MTB SigA	(57)
pMTBRP-5 thru 8	Original constructs for MTB RNAP Expression	(56)
pACYCDuet-rpoA-rpoZ	Dual expression vector for MTB RNAP $\alpha$ and $\omega$ subunits	This study
pETDuet-rpoB-rpoC	Dual expression vector for MTB RNAP $\beta$ and $\beta'$ subunits	This study
pRSF-SigA	Expression vector for MTB SigA	This study

### Purification of MTB RNAPs

MTB RNAP holoenzymes (WT, D435V, H445Y, S450L, and S450L/V483G) were purified as follows. *E. coli* BL21(DE3) containing the expression vectors pETDuet-rpoB-rpoC, pACYCDuet-rpoA-rpoZ, and pRSF-SigA were grown at 37°C in 1 L 2xTY with carbenicillin, kanamycin, and chloramphenicol until OD<sub>600</sub> = 0.4 – 0.6. Protein expression was induced with 0.5 mM IPTG and grown for 16 to 20 hrs at 16°C. Cells were harvested, and the protein was purified and stored as described in Chapter II for *E. coli* WT RNAP, using TGED containing 300 mM NaCl for gel filtration with the Sephacryl S-300 HR column.

### Purification of MTB SigA

*E. coli* BL21(DE3) containing pMCSG7-SigA were grown in 2 L 2xTY at 37°C until OD<sub>600</sub> = 0.4 – 0.6, then expression was induced with 1 mM IPTG and grown at 16°C for 16 – 20 hrs. Cells were pelleted and resuspended in  $\sigma^{70}$  lysis buffer and supplemented with 2 mM PMSF. Cells

were sonicated and lysate was cleared by centrifugation (21,000 x g, 40 min, 4°C). Supernatant was incubated with 2 mL Ni-NTA agarose and incubated at 4°C with shaking for 1 hr. Mixture was then applied to a column and washed as follows: 5 CVs –  $\sigma^{70}$  lysis buffer, 8 CVs –  $\sigma^{70}$  lysis buffer + 15 mM imidazole, 10 CVs –  $\sigma^{70}$  lysis buffer + 100 mM imidazole. Eluted protein was concentrated using a 10 kDa MWCO Amicon Ultra-15 Centrifugal Filter Unit, filtered, and applied to a HiPrep 16/60 Sephacryl S-200 HR equilibrated with TGED containing 200 mM NaCl. Protein was pooled and concentrated, then flash-frozen and stored at -80°C.

### **Preparation of Materials for High-Throughput Screen and Transcription Assays**

Large quantities of pMGA12 were generated for the screen using a Qiagen Giga Kit, preparing Buffer P1 without RNase A and regenerating the QIAGEN tip-10000 columns as previously described.<sup>92</sup> NTPs were prepared at 100 mM, then adjusted the pH of each NTP to 7.5 before preparing a 25 mM NTP mix by adding equal amounts of each NTP. Buffers used for the screen were prepared with certified DNase/RNase free components and Milli-Q water.

### **Time Course Transcription Assay**

Reactions of 100  $\mu$ L for each *E. coli* RNAP were prepared as follows. In each tube, 48  $\mu$ L of a 2x *E. coli* reaction buffer solution [80 mM Tris-HCl (pH = 7.5 at 23°C), 100 mM KCl, 20 mM MgCl<sub>2</sub>, 0.02% Triton X-100, 16 mM DTT, 20 nM pMGA12, 1 mM NTPs] was added, 2  $\mu$ L of DMSO, followed by 48  $\mu$ L of a 2x RNAP solution (80 nM RNAP, 240 nM  $\sigma^{70}$ ). Reactions were incubated at 37°C for 3.5 hrs. Starting at time = 0 min, 10  $\mu$ L aliquots were removed from the reaction mixture every 30 min and quenched in 40  $\mu$ L ice cold Milli-Q water in a 96-well plate (Corning 96-Well Half Area Black Flat Bottom Polystyrene NBS™ microplate). Then, 50  $\mu$ L of 150  $\mu$ M MG in water was added to each well and incubated on ice for 10 min before measuring fluorescence at

excitation and emission wavelengths of 628 and 660 nm, respectively. Raw fluorescence values were plotted against time for each enzyme tested.

MTB RNAPs were tested in the same manner, using a 2x MTB reaction buffer solution (40 mM Tris-HCl (pH = 8.0 at 37°C), 20 mM MgCl<sub>2</sub>, 300 mM potassium L-glutamate, 15% glycerol, 20 nM pMGA12, 1 mM NTPs), and used MTB SigA instead of *E. coli*  $\sigma^{70}$ .

### **High-throughput Bacterial Transcription Assay**

The transcription assay described above was modified to 384-well plate format, with a reaction volume of 20  $\mu$ L. A 2x *E. coli* reaction buffer was added (10  $\mu$ L) to each well of 384-shallow well microplates (ProxiPlate-384 Plus F) using a multidrop dispenser (Thermo Scientific, Waltham, MA). Compounds or DMSO were then added (20  $\mu$ M final, n = 1, 0.2  $\mu$ L) to the appropriate wells using a Biomek FX HDR pintool instrument (Beckman, Fullerton, CA). TGED containing 50 mM NaCl and 0.01% Triton X-100 was added (10  $\mu$ L) to the appropriate wells for a positive control via multidrop dispenser. A 2x RNAP solution (80 nM  $\beta$ S531L/ $\beta$ 'V408G *E. coli* RNAP core, 240 nM *E. coli*  $\sigma^{70}$  in TGED containing 50 mM NaCl and 0.01% Triton X-100 was added to the appropriate wells (10  $\mu$ L) via multidrop dispenser for a reaction volume of 20  $\mu$ L. The plates were incubated at 37°C for 2 hrs, then placed on ice for 10 min. A solution of 150  $\mu$ M malachite green was added (5  $\mu$ L) to each well. The plates were incubated on ice for 10 min before measuring fluorescence with an EnVision Multilabel plate reader (PerkinElmer) using a Cy5 620/10 filter for excitation and a LANCE 665/7.5 filter for emission.

### **Reconfirmation and Counter Screens**

Compounds identified from the primary HTS were tested against a series of control screens. First, compounds were retested (n = 3) in the assay following the protocol above, except

compounds were added to empty plates using a *mosquito X1* (TTP Labtech) prior to addition of buffer and RNAP solutions. These compounds were also tested for interference of MG-MGA fluorescence by adding 20  $\mu\text{L}$  of a reaction mixture (1x *E. coli* reaction buffer solution and 1x *E. coli* RNAP solution incubated at 37°C for 2 hrs) into plates containing the compounds, followed by addition of malachite green (5  $\mu\text{L}$ , 150  $\mu\text{M}$ ). After 10 min on ice, the fluorescence was measured at excitation and emission wavelengths of 620 and 665 nm, respectively (EnVision plate reader). Compounds that reconfirmed and did not inhibit MG-MGA fluorescence were tested for concentration-dependent inhibition. For the concentration-response study (n=2), the concentration of the compounds was varied using two-fold serial dilutions (0.78 – 100  $\mu\text{M}$  final), and tested following the screening protocol described above, adding the compounds to empty plates using the *mosquito X1*.

### **Data Analysis**

Active compounds from the initial high-throughput screen were defined as compounds that decreased fluorescence signal more than or equal to three times the standard deviation (s.d.) of the negative controls (all analyses were calculated on a plate-by-plate basis) and resulted in greater than or equal to 20% fluorescence inhibition. Compounds that the NIH does not allow in their screening collection due to reactivity or known toxic substituents (black flags), and compounds with chemical motifs that tend to be highly reactive or highly toxic and nonspecific (red flags) were excluded.<sup>93</sup> Compounds that hit in more than 30% of screens previously run at the CCG were also excluded.

Compounds that passed these initial selection criteria were retested (n=3) and screened for false positives due to inhibition of fluorescence of MGA bound malachite green. Compounds

that inhibited fluorescence of MGA-bound malachite green (n=3) more than 25% were excluded. (Note that two rows on each plate used a less stringent cutoff of 45% fluorescence inhibition due to a partial clog in one tube of the multidrop dispenser cassette for MG, resulting in a lower average fluorescence for these wells.) Remaining compounds with fluorescence inhibition more than or equal to three times the s.d. of the negative controls in at least 2 wells (n=3) were tested for concentration-dependent inhibition. Compounds that inhibited *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP in a concentration-dependent manner (determinable IC<sub>50</sub> or displayed concentration-dependent trend) were considered hits from the high-throughput screen. Data were fit by nonlinear regression to Eq. (1) where M0 = log of compound concentration, M1 = log of IC<sub>50</sub>, M2 = hill slope, M3 = lower limit of the curve, and M4 = upper limit of the curve.

$$(1) \quad y = M3 + \frac{(M4 - M3)}{(1 + 10^{(M0 - M1) \times M2})}$$

### Reconfirmation of Hits

Fresh samples of compounds identified as hits were ordered from commercial vendors for a reconfirmation concentration-response study. This study was done in the same format as the initial concentration-response study, testing the fresh compounds for activity against *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP. The fresh compounds were then tested against a panel of *E. coli* and MTB RNAPs (10 total, Table III-3). The compounds were tested against each enzyme using the screening protocol above (using 2x MTB reaction buffer for MTB RNAPs). Two-fold serial dilutions of compounds were dispensed in duplicate from a mother plate using the pintool device. The mother plate concentrations ranged from 20 to 0.01 mM, giving final compound concentrations of 200 to 0.1  $\mu$ M in the transcription reactions. Rifampin was tested in the same manner, using



a mother plate to give final compound concentrations of 1280 to 0.625 nM in the transcription reactions.

## Results

### Plasmid Preparation and Expression of *E. coli* and MTB RNAPs

To evaluate the inhibition of compounds identified from the HTS against a panel of bacterial RNAPs (*E. coli* and MTB, WT and Rif<sup>R</sup>), plasmid constructs for *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP and MTB RNAPs were prepared to allow for the expression and purification of each RNAP of interest. A pVS10 construct previously prepared for expression of *E. coli*  $\beta$ S531L RNAP was modified by mutagenesis of the *rpoC* gene to encode the compensatory mutation V408G, resulting in the mutant vector pVS10(S531L/V408G).<sup>56</sup> The expression system for MTB RNAP was prepared as described by Banerjee et al. to give the plasmid vectors pACYCYDuet-rpoA-rpoZ and pETDuet-rpoB-rpoC.<sup>91</sup> The mutated genes for the MTB RNAPs D435V, H445Y, and S450L were subcloned individually to give mutant vectors for each on the pETDuet-rpoB-rpoC plasmid construct. The compensatory mutation  $\beta$ 'V483G was subcloned onto pETDuet-rpoB(S450L)-rpoC for preparation of the expression vector pETDuet-rpoB(S450L)-rpoC(V483G). The *E. coli* and MTB RNAPs were successfully purified and presence of subunits was verified by SDS-PAGE. The concentrations of the proteins were determined by the Bradford protein assay, and the activity of the RNAPs were confirmed using the plasmid-based transcription assay (data not shown).

### Time Course Transcription Assay

The linearity of transcription as a function of time was tested for each enzyme purified to verify the length of time for the end-point detection transcription assay. It was important to ensure that each enzyme activity was still linear after the 2-hr detection time used in the

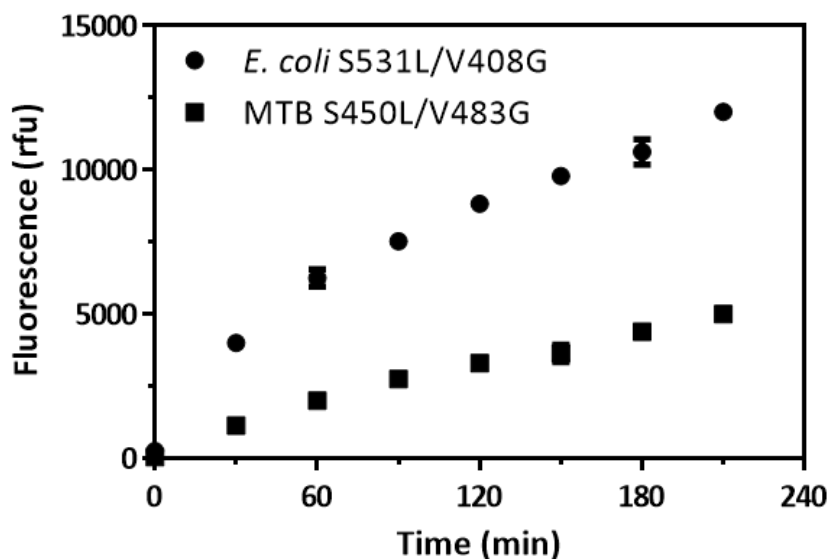


Figure III-1. Transcription activity of *E. coli* and MTB Rif<sup>R</sup> RNAP with compensatory mutation. *E. coli*  $\beta$ S531L/ $\beta$ V408G and MTB  $\beta$ S450L/ $\beta$ V483G RNAPs were tested in the plasmid-based transcription assay, with 10  $\mu$ L aliquots being quenched every 30 min. Fluorescence detection by malachite green (Ex/Em = 628/660 nm).

transcription assay in order to be able to accurately determine and compare inhibition between compounds and across enzymes. The transcription activity of *E. coli*  $\beta$ S531L/ $\beta$ V408G RNAP used for the HTS and its MTB homolog are shown in Figure III-1.

It was determined that the RNAP transcription activity was linear beyond 2 hrs (tested to 3 hrs for each enzyme) for each enzyme (data only shown for one representative enzyme for each species, *E. coli* and MTB). This indicates that the two-hour reaction for end-point detection is sufficient when using the plasmid-based transcription assay to detect RNAP activity and the effect of compounds on enzyme activity.

### Adaptation for High-Throughput Screen

The plasmid-based transcription assay was adapted for an HTS by decreasing the reaction size from 50  $\mu$ L to a 20  $\mu$ L reaction. The transcription reaction was tested in various 384-well plates (Corning 4514, Greiner 784076, Greiner 784900, and ProxiPlate-384 Plus F) to determine

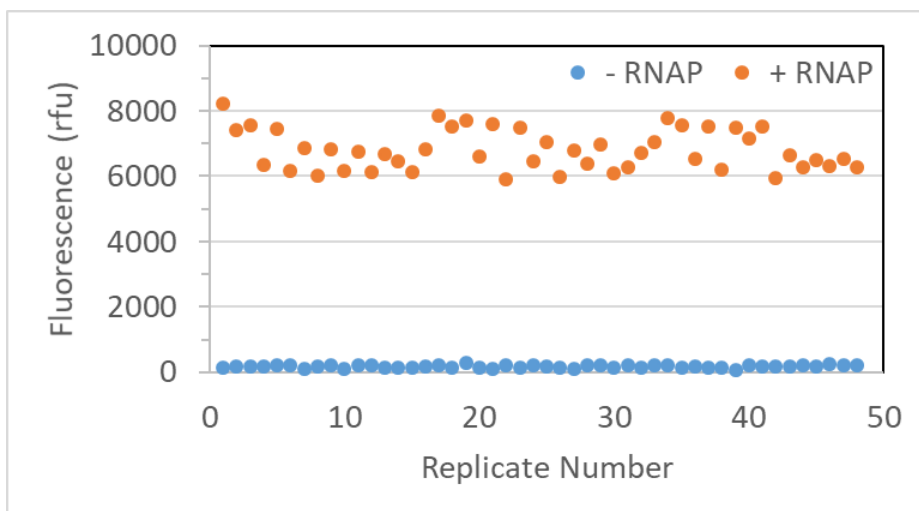


Figure III-2. Controls of transcription assay with *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP in 384-well plate format for high-throughput screen. Positive control values (no RNAP) are at the bottom in blue, and negative control values (RNAP, DMSO) are at the top in orange.  $Z'$  was determined using Eq. (2).

the optimal plate for screening (data not shown). It was determined that the ProxiPlate resulted in the largest signal for transcription detection, and was therefore chosen as the assay plate for the HTS. The transcription assay was tested at the high-throughput screening facility using all the liquid handling equipment, testing the positive and negative controls (no RNAP and DMSO, respectively). An assay plate with *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP containing the negative and positive controls at a ratio of 50:50 were tested to determine the suitability of the assay for an HTS (Figure III-2). The suitability of the assay for an HTS was determined by calculating the Z-factor ( $Z'$ ) of the assay when conducted using the liquid handling equipment. The  $Z'$  is a dimensionless, statistical characteristic for an HTS assay that can be utilized for comparison and evaluation of the quality of an assay.<sup>94</sup> It is calculated from the means and s.d. of both the positive and negative controls of the assay, using Eq. (2) where  $\sigma_p$  and  $\sigma_n$  represent the s.d. of the positive and negative control (respectively) and  $\mu_p$  and  $\mu_n$  represent the mean of the positive and negative control (respectively).

$$(2) \quad Z' = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

A  $Z' > 0.5$  is considered suitable for an HTS, whereas  $Z' < 0.5$  can have an increased propensity for false positives and false negatives. False positives and negatives are compounds that are incorrectly assigned as active or inactive. This can be due to assay interference, detection interference, or the uncertainty/error associated with the assay. The  $Z'$  for the assay using the test plate in Figure III-2 was determined to be 0.70, indicating that the transcription assay is highly suitable for an HTS.

### High-Throughput Screen

The ~150,000 compound HTS was carried out using the *E. coli*  $\beta$ S531L/ $\beta'$ V408G RNAP with pMGA12. The average  $Z'$  per plate in the primary screen was calculated to be 0.86 across the plates tested in the screen. Figure III-3 is an inhibition plot for the primary screen, showing the percent inhibition values (calculated on a plate-by-plate basis) for each of the compounds tested. The 20% inhibition line (Figure III-3, blue) was one of the selection criteria for the initial HTS triage.

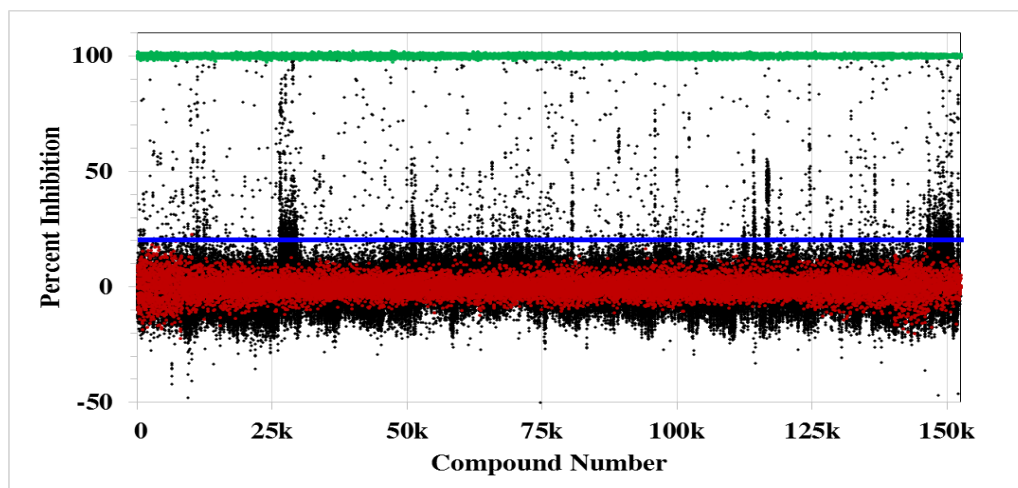


Figure III-3. Inhibition plot of 150,554 compounds tested in the HTS against *E. coli*  $\beta$ S531L/ $\beta'$ V408G RNAP. Compound number is an arbitrary reassignment of CCG compound numbers starting at 1 for visualization purposes. Positive control values are green, negative control values are red, and sample values are black. Blue line indicates 20% inhibition.

## HTS Hit Identification

A summary of the triage of the HTS is shown in Figure III-4. The primary screen identified 1738 compounds (1.2% hit rate) that met the initial selection criteria of percent inhibition, promiscuity, and potential toxicity and reactivity. The 1738 compounds were then tested in two different control screens. The compounds were retested in triplicate and tested for inhibition of the MG-MGA fluorescence signal (in triplicate). This resulted in the identification of 307 compounds that reconfirmed in triplicate and did not inhibit the MG-MGA fluorescence signal. These 307 compounds were then tested for concentration-dependent inhibition, which identified 234 compounds that displayed concentration-dependent inhibition of *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP transcription, with 190 compounds having a determinable IC<sub>50</sub> (IC<sub>50</sub> < 100  $\mu$ M, the highest

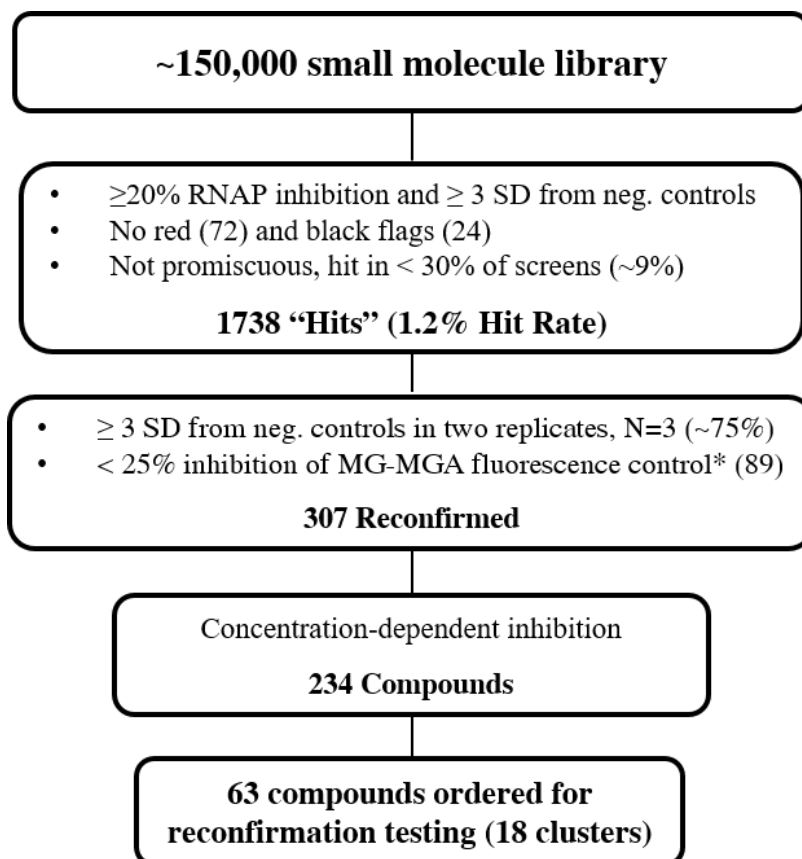


Figure III-4. Flow chart of selection criteria for high-throughput screen. \*A less stringent cutoff (45%) was used for two rows of the 384-well plates (see Results: Data Analysis).

concentration tested). Figure III-5 contains a waterfall plot of the IC<sub>50</sub> values for the compounds that showed concentration-dependent inhibition. With the help of Drs. Paul Kirchhoff and Hollis Showalter, the compounds were further triaged based on their potential toxicity, novelty, redundancy, and commercial availability. This resulted in the selection of 63 compounds that fell into 18 different structural scaffold clusters to be ordered.

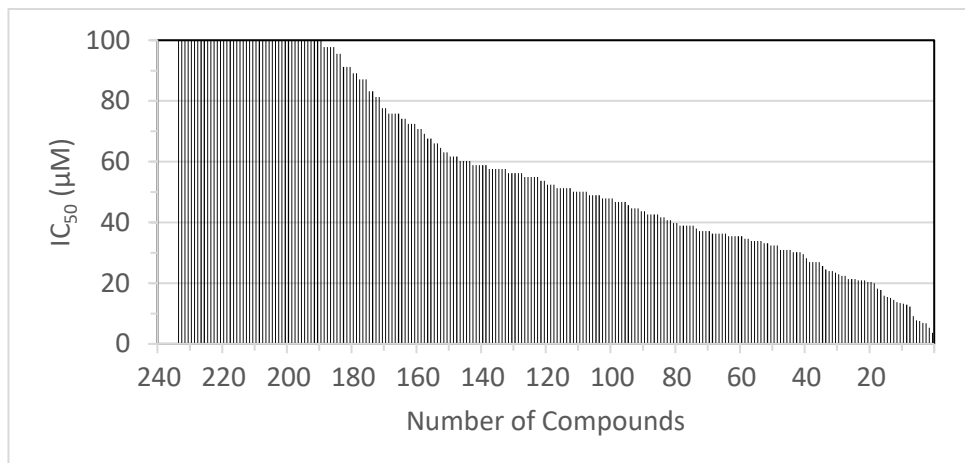


Figure III-5. Waterfall plot of IC<sub>50</sub> values determined against *E. coli* βS531L/β'V408G of 234 compounds identified during HTS triage.

### Reconfirmation of Hit Compounds

Rifampin was tested for inhibition of *E. coli* WT RNAP using the liquid handling equipment to validate the accuracy of the instrumentation, particularly the Biomek FX HDR pintoole instrument, since using a pintoole is an atypical way to set up concentration-response assays. In parallel, transcription reactions testing inhibition of *E. coli* WT RNAP by rifampin in a concentration-dependent manner were setup either 1) by hand using multipipettes or 2) using the HTS liquid handling equipment. Concentration-response curves for the two experiments were analyzed by nonlinear regression. As seen in Figure III-6, the two experimental methods resulted in similar concentration-response curves, indicating the accuracy of the pintoole when setting up concentration-response curves.

### Rifampin Inhibition of *E. coli* WT RNAP

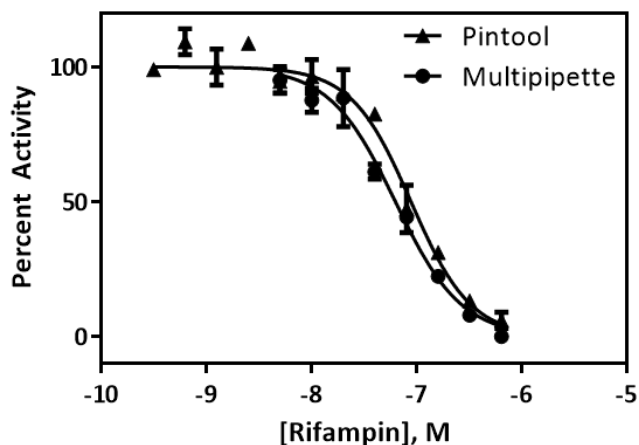


Figure III-6. Comparison of compound transfer methods for transcription assay inhibition studies. Biomek FX HDR pintooler was used to transfer 0.2  $\mu$ L of varying concentrations of rifampin in DMSO.

Fresh samples of the 63 compounds were purchased and re-tested for activity. Of the 63 compounds, 30 compounds were inactive against *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP; however, only three were eliminated due to the other 27 inactive compounds being members of structural clusters that contained active analogs. The 60 compounds were then tested for activity against a panel of *E. coli* and the corresponding MTB RNAPs: WT, 3 different Rif<sup>R</sup> mutants ( $\beta$ D516V,  $\beta$ H526Y,  $\beta$ S531L – *E. coli* numbering), as well as the MTB double mutant with a compensatory mutation ( $\beta$ S450L/ $\beta$ 'V483G). The IC<sub>50</sub> values of the compounds against these RNAPs ranged from 5  $\mu$ M to >200  $\mu$ M, and most of the compounds had similar values across the RNAP enzymes tested. The individual IC<sub>50</sub> values for the most active compounds against each enzyme are in Table III-3.

### Discussion

*M. tuberculosis* RNAP is a well-validated target for the treatment of tuberculosis; however, target-based mutations conferring resistance to rifampin have resulted in a large number of Rif<sup>R</sup> strains in clinical isolates. The potent sterilizing activity (i.e., bactericidal against

both active and latent MTB) of rifampin against WT MTB indicates that RNAP is still a promising target for the treatment of Rif<sup>R</sup> MTB if the rifamycin resistance can be circumvented. An assay that monitors the increase in fluorescence of malachite green upon binding to an RNA aptamer has been adapted for high-throughput screening. Obtaining quantities of *M. tuberculosis* RNAP sufficient for a high-throughput screen has been prohibitive. However, the *E. coli* RNAP can be obtained in sufficient quantities, has greater activity, and is more stable than the MTB RNAP. We have shown that the *E. coli* RNAP is a reasonable model for the MTB RNAP<sup>56</sup> and we have solved the X-ray crystal structure of the *E. coli* RNAP alone and in complex with rifamycin derivatives that we have designed.<sup>95,96</sup> This provides us the potential to determine the three-dimensional structure of the RNAP complex with any novel RNAP inhibitor that we find which has sufficient affinity.

An HTS of 150,554 small molecules was carried out against *E. coli* RNAP  $\beta$ S531L/ $\beta$ 'V408G (Figure III-3) using the plasmid-based *in vitro* transcription assay. As shown in Figure III-4, the concentration-response, control screening, and triage for potential toxicity, novelty, redundancy, and commercial availability resulted in the selection of 63 compounds to order.

The reconfirmation assays were setup using the Biomek pintooleer for multiple reasons. First, it can dispense small volumes of compound quickly from a stock plate to an assay plate. Second, unlike the *mosquito X1*, the Biomek pintooleer can transfer compounds very quickly (minutes instead of hours) and can easily be used to set up multiple plates containing the same set of compound dilutions. As was shown with rifampin against the *E. coli* WT RNAP, the pintooleer can accurately dispense compound dilutions and replicate the inhibition obtained through traditional pipette transfer techniques (Figure III-6).



Table III-3. In vitro activity of selected compounds from HTS against *E. coli* and MTB RNAPs

Compound	Cluster	<i>E. coli</i> RNAP, IC <sub>50</sub> (μM)						MTB RNAP, IC <sub>50</sub> (μM)						Vendor (Purity)
		WT	D516V	H526Y	S531L	S/V <sup>a</sup>	WT	D435V	H445Y	S450L	S/V <sup>a</sup>			
257176	1	61	33	42	40	22	43	31	20	42	72	Enamine (90%)		
257180	1	8.4	7.1	9.1	6.7	11	4.8	4.6	3.7	5.7	8.9	ChemDiv (90%)		
257181	1	22	14	20	18	18	21	14	11	19	22	ChemDiv (90%)		
257182	1	13	10	13	10	13	12	7.7	6.8	10	14	ChemDiv (90%)		
19247	2	21	22	36	17	22	12	6.6	4.5	12	27	Vitas-M (92%)		
23486	3	28	26	25	19	19	19	14	7.7	16	32	Vitas-M (92%)		
257178	3	16	16	17	18	13	12	7.8	8.3	13	15	ChemDiv (90%)		
123513	4	11	5.9	7.7	5.9	26	6.3	2.2	2.6	8.7	20	Vitas-M (92%)		
257177	4	190	110	120	140	47	55	32	18	47	39	Pharmeks (90%)		
22785	5	48	33	45	44	30	53	14	16	42	67	Vitas-M (92%)		
170440	6	6.5	12	(25)	14	91	(38)	(45)	150	190	(39)	Vitas-M (92%)		
170444	6	46	66	NA	29	-43	(40)	120	100	(49)	(49)	Vitas-M (92%)		
186618	6	11	10	99	15	39	28	11	21	15	13	ChemDiv (90%)		
193157	6	18	18	(35)	47	46	(33)	(29)	(46)	(33)	(44)	Vitas-M (92%)		
145060	7	NA	NA	NA	NA	(41)	64	17	12	61	110	Enamine (90%)		

<sup>a</sup>S/V = *E. coli* βS531L/β<sup>v</sup>V408G or MTB βS450L/β<sup>v</sup>V483G

NA = No inhibition at highest concentration tested

Values in parentheses are % inhibition at highest concentration, 200 μM (120 μM for *E. coli* S/V)

Error for IC<sub>50</sub> values are 10 – 25%

Table III-4. Compounds with similar activity across WT and Rif<sup>R</sup> RNAP

Compound Name	Structure	Cluster	Average IC <sub>50</sub> (μM)	
			<i>E. coli</i>	MTB
<b>Compounds with Similar Activity Between <i>E. coli</i> and MTB RNAPs</b>				
257176		1	40 ± 14	41 ± 19
257180		1	8 ± 2	6 ± 2
257181		1	18 ± 3	17 ± 5
257182		1	12 ± 2	10 ± 3
19247		2	23 ± 7	12 ± 9
23486		3	24 ± 4	18 ± 9
257178		3	16 ± 2	11 ± 3
123513		4	11 ± 8	8 ± 7
22785		5	40 ± 8	38 ± 23
<b>Compounds More Active Against MTB RNAP</b>				
257177		4	140 ± 36	38 ± 14
145060		7	>200	53 ± 40

IC<sub>50</sub> values are averages of WT, S531L, D516V, H526Y, βS531L/β'V408G (*E. coli* numbering). See Table III-3 for individual IC<sub>50</sub> values.

Table III-5. Compounds with reduced activity against *E. coli* H526Y RNAP (cluster 6)

Compound Name	Structure	<i>E. coli</i> RNAP					MTB
		WT	D516V	H526Y	S531L	S/V <sup>a</sup>	
186618		11	10	99	15	39	17 ± 7 <sup>b</sup>
170440		6.5	12	>200	14	91	>150
170444		46	66	>200	29	>120	>100
193157		18	18.2	>200	46.9	46	>200

<sup>a</sup>*E. coli* βS531L/β'V408G RNAP

<sup>b</sup>IC<sub>50</sub> values are averages of WT, S450L, D435V, H445Y, βS450L/β'V483G. See Table III-3 for individual MTB IC<sub>50</sub> values.

The fresh samples were screened against a panel of *E. coli* and the corresponding MTB RNAPs: wild-type, 3 different Rif<sup>R</sup> mutants (D516V, H526Y, S531L – *E. coli* numbering) as well as the double mutant ( $\beta$ S531L/ $\beta$ 'V408G – *E. coli* numbering). The IC<sub>50</sub> values of the compounds against these RNAPs ranged from 5 to > 200  $\mu$ M (data not shown). With the exception of cluster 6, discussed below, we selected scaffolds that do not show any large (>10-fold) differential in IC<sub>50</sub> values across the panel of WT and Rif<sup>R</sup> RNAP mutants tested, as these would seem to be more promising as they are likely to be binding outside of the rifampin resistance determining region, hence lacking cross-resistance with Rif<sup>R</sup> RNAPs.

Encouragingly, 9 of the 63 compounds (Table III-4) had similar IC<sub>50</sub> values across the panel of RNAP enzymes tested, indicating that they are active against Rif<sup>R</sup> mutants. Cluster 1, the cyanopyrimidine scaffold was identified as one of the more potent scaffolds, with IC<sub>50</sub> values ranging from 5 to 40  $\mu$ M.

Another group of compounds (Table III-5, cluster 6, pyrazolopyrimidines) showed comparable activity across most of the *E. coli* RNAPs with the exception of H526Y. A recent crystal structure of *E. coli* H526Y RNAP shows significant conformational changes in the rifamycin binding pocket relative to the wild-type, D516V, and S531L RNAPs.<sup>97</sup> This suggests that these compounds may bind in the rifamycin binding pocket, but their much smaller size relative to rifampin allows them to bind to the D516V and S531L Rif<sup>R</sup> mutants. As precedence, the RNAP inhibitor sorangicin A has been shown to bind in the rifamycin binding pocket and encounter partial resistance by some Rif<sup>R</sup> mutants and complete cross-resistance by other Rif<sup>R</sup> mutants.<sup>62,98</sup> Xu et al. showed that *E. coli* H526Y RNAP had complete resistance to sorangicin A while mutants D516N and S531F, and other residues, showed only partial or no resistance.<sup>99</sup> Compound 186618

of cluster 1 was however active against all of the RNAP enzymes tested, indicating a potential for activity against *E. coli* H526Y RNAP and MTB RNAPs.

Two compounds that are more potent against MTB RNAPs relative to *E. coli* RNAP (Table III-4) were identified. Compound 257177 is in the same cluster as 123513 (Table III-4, cluster 4), but has a significantly different activity profile. Compound 257177 is ~10-fold less active against *E. coli* RNAP than 123513. We cannot make any significant correlations between the structures and activities of these two compounds; however, these differences are noteworthy for potential future structure activity relationship studies in optimizing activity for MTB (or *E. coli*) RNAP. Compound 145060 showed no significant activity against *E. coli* RNAP in the follow up assays, and only modest activity against MTB RNAPs (>50  $\mu$ M).

Eleven of the fifteen compounds (Table III-4) exhibited very similar activity against WT and Rif<sup>R</sup> RNAPs, providing proof of principle that novel RNAP inhibitors can be found that are insensitive to the Rif<sup>R</sup> mutations. While the in vitro activities of these analogs are very modest (low micromolar at best), these compounds represent an initial structure-activity relationship (SAR) probe using only commercially available HTS hits. A hit-to-lead campaign involving iterative design, synthesis and testing will very likely identify much more potent inhibitors of WT and Rif<sup>R</sup> MTB RNAP.

## Conclusions

We have adapted a plasmid-based transcription assay for use in a high-throughput screen, and screened 150,000 compounds against a Rif<sup>R</sup> bacterial RNAP (*E. coli*  $\beta$ S531L/ $\beta$ 'V408G). The high-throughput screen identified 234 compounds that inhibited *E. coli*  $\beta$ S531L/ $\beta$ 'V408G RNAP in a concentration-dependent manner. After further triage, we purchased 63 compounds

for further testing. The purchased compounds were tested against both WT and Rif<sup>R</sup> mutants (D516V, H526Y, S531L – *E. coli* numbering) of *E. coli* and MTB RNAPs as well as MTB  $\beta$ S450L/ $\beta'$ V483G RNAP. We have identified 15 compounds of interest that were active against either *E. coli* or MTB RNAPs, or both.

## Notes to Chapter III

### Acknowledgements

I would like to thank Dr. Suman Gill for her work in preparation of the pMTBRP constructs and pVS10 constructs used in this study. I would like to thank Dr. Katsu Murakami (Penn State University) for providing the expression vector pET15b-Sig70. Dr. Irosha Nawarathne for construction of pVS10(S531L/V408G). I would like to thank Maxwell Stefan for his work in preparation and optimization of the MTB expression system that was used for the MTB RNAP purifications. I would like to greatly thank Dr. Vadim Molodtsov (Penn State University) for preparation of the RNAP and  $\sigma^{70}$  used for the HTS. I would like to thank Martha Larsen for her expertise on high-throughput screening and valuable input when implementing the assay for high-throughput screening. Dr. Nicholas Santoro and Tom McQuade at the Center for Chemical Genomics (CCG) for their help with the HTS and sample management. Drs. Paul Kirchhoff and Hollis Showalter for their expertise and help in the HTS triage process.

The work described in this chapter was published in *SLAS Discovery* in 2017 (Volume 22, Issue 3; pages 287 – 297).

### Abbreviations

$\beta$ -ME,  $\beta$ -mercaptoethanol; CCG, Center for Chemical Genomics; CV, column volume; DMSO, dimethylsulfoxide; DNA, deoxyribonucleic acid; DTT, dithiothreitol; EDTA, ethylenediamine tetraacetic acid; HTS, high-throughput screen;  $IC_{50}$ , half-maximal inhibitory concentration; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; MDR-TB, multidrug-resistant tuberculosis; MG, malachite green; MGA, malachite green aptamer; MTB, *Mycobacterium tuberculosis*; MWCO, molecular weight cutoff; NTP, nucleoside triphosphate;  $OD_{600}$ , optical density at 600 nm; PAINS, pan-assay interference compounds; PMSF, phenylmethylsulfonyl fluoride; Rif<sup>R</sup>, rifampin-resistant; RNA, ribonucleic acid; RNAP, RNA polymerase; SAR, structure-activity relationship; s.d., standard deviation; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TB, tuberculosis; WT, wild-type; XDR-TB, extensively drug-resistant tuberculosis

## Chapter IV

### In vitro Evaluation of Hit Scaffolds Identified from High-Throughput Screen

Bacterial RNA polymerase (RNAP) is an antibacterial target of interest for several reasons. Bacterial RNAP is an essential enzyme that has been an important target for the treatment of tuberculosis (TB) since the introduction of rifamycins in the 1960s. The primary mechanism of bacterial resistance arises from site-specific mutations in the rifampin binding site. It has already been shown that potent compounds can be identified that do not have cross-resistance to rifampin, as seen with GE23077.<sup>63</sup> Therefore, in order to identify a potent small molecule inhibitor of wild-type (WT) and rifampin-resistant (Rif<sup>R</sup>) *Mycobacterium tuberculosis* (MTB) RNAP, a high-throughput screen (HTS) was performed against the Rif<sup>R</sup> *E. coli* RNAP homolog  $\beta$ S531L/ $\beta$ 'V408G RNAP. Initial triage resulted in the purchase and evaluation of 63 compounds identified from the HTS, identifying 15 compounds with promising activity against a panel of *E. coli* and/or MTB RNAPs.

Scaffolds identified from the HTS (Chapter III) were further investigated for their potential as potent inhibitors of bacterial RNAP. Since the compounds identified from the screen have low activity (low micromolar at best), the scaffolds were first evaluated for potential structure-activity relationships (SAR) using commercially available analogs. This was done to determine if there are changes in their activity correlated with changes in the compound structure – which would indicate that there is a distinct binding site and that binding affinity and inhibition may be



improved in a hit-to-lead campaign. In addition, select compounds of interest were synthesized to further validate the structure and activity of the purchased compounds.

Aromatic nitriles have been reported to form covalent bonds with enzymatic cysteines and lead to irreversible covalent inhibition depending on their electrophilicity (Figure IV-1).<sup>100</sup> Inhibitors that covalently interact with a cysteine have been developed for targets such as kinases with clinical success.<sup>101–103</sup> However, there are potential issues associated with covalent inhibitors, such as non-specificity.<sup>104</sup> In order to determine if the cyanopyrimidine scaffold that we identified in our HTS is a false positive and non-specifically interacting with bacterial RNAP, the scaffold was investigated for 1) essentiality of nitrile for inhibition, 2) reversibility of inhibition, and 3) cysteine adduct formation. The results of these experiments were examined to determine whether the cyanopyrimidine scaffold is a specific inhibitor of bacterial RNAP and a potential candidate for further optimization as an anti-TB agent.

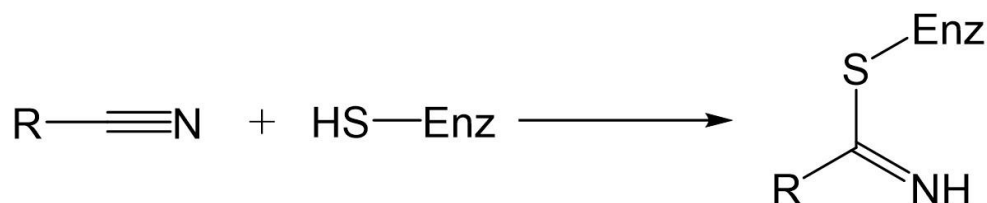


Figure IV-1. Covalent interaction of nitrile with L-cysteine to form cysteine adduct.

## Materials and Methods

### Materials

All of the materials were the same as those specified in Chapter II. Compound analogs purchased were from ChemDiv or Vitas-M, unless otherwise specified. Synthetic oligonucleotides purchased from Invitrogen are listed in Table IV-1. A Shimadzu LC-MS was used

for nitrile reactivity assay analysis. pMGA4-Ec and pMGA4-Mtb were prepared as described for pMGA12 in Chapter III. *E. coli* and MTB RNAPs were prepared as described in Chapter III.

Table IV-1. List of oligonucleotides used for FID Assay

Oligonucleotide Name	Sequence, 5' - 3'	Length
FID-lacUV5 Top	TTAGGCACCCCAGGCTTTACACTTTATGCTTCC GGCTCGTATAATGTGTGGAATTGTGAG	60
FID-lacUV5 Bottom	CTCACAATTCCACACATTATACGAGCCGGAAG CATAAAGTGTAAGCCTGGGGTGCCTAA	60

### Fluorescent Intercalator Displacement Assay

Binding of compounds to DNA was determined by fluorescent intercalator displacement (FID). A 60-bp FID-lacUV5 probe was used for this study. The FID-lacUV5 probe was hybridized by mixing the top and bottom oligonucleotides (Table IV-1) 1:1 after suspending in TE/NaCl buffer (10 mM Tris-HCl, 1 mM EDTA, 50 mM NaCl, pH = 8.0). The top and bottom oligonucleotides were annealed in a thermocycler by heating at 94°C for 2 min, followed by 70°C for 2 min, then was allowed to slowly cool to room temperature. The assay was conducted in triplicate as previously described.<sup>105</sup> In brief, 70  $\mu$ L of 64.3  $\mu$ M ethidium bromide in Tris/NaCl buffer (0.1 M Tris HCl, 0.1 M NaCl, pH = 8.0) was added to all wells of a black, half-area, 96-well plate. Then, 10  $\mu$ L of TE/NaCl buffer (background fluorescence control) or 15  $\mu$ M FID-lacUV5 probe in TE/NaCl buffer were added to the appropriate wells, followed by addition of 20  $\mu$ L test compound in 10% DMSO Tris/NaCl Buffer (10% DMSO Tris/NaCl Buffer for 100% fluorescence control). The concentrations tested ranged from 12.5 to 100  $\mu$ M. The plates were protected from light and allowed to equilibrate for 30 min followed by fluorescence measurement (Ex/Em = 545/595 nm). The background fluorescence was subtracted from all data and % fluorescence inhibition was

calculated according to Eq. (1) where  $F_{\text{test}}$  = average fluorescence obtained in presence of test compound and  $F_{100\%}$  = the average fluorescence of the 100% fluorescence control).

$$(1) \quad \% \text{ Fluorescence Inhibition} = \frac{F_{100\%} - F_{\text{test}}}{F_{100\%}} \times 100$$

### **Minimal Inhibitory Concentrations against virulent *M. tuberculosis* and Cytotoxicity**

Minimal inhibitory concentrations (MICs) of compounds against *M. tuberculosis* H<sub>37</sub>R<sub>V</sub> were determined by Baojie Wan at the University of Illinois at Chicago as previously described using the Microplate Alamar Blue Assay (MABA) and Low-Oxygen Recovery Assay (LORA).<sup>106</sup> Cytotoxicity of compounds were determined by Baojie Wan against Vero cells using an MTS assay as previously described.<sup>107</sup>

### **Plasmid-Based Transcription Assay**

Plasmid-based in vitro transcription assays to determine compound IC<sub>50</sub> values were performed as described in Chapter II for rifampin with the following modifications. The concentration of pMGA12 or pMGA4-Ec (Figure IV-2) was 80 nM in the 2x *E. coli* reaction buffer, and *E. coli* WT RNAP and  $\sigma^{70}$  were 40 nM and 160 nM (respectively) in the 2x RNAP solution. Reactions were incubated at 37°C for 90 min before quenching in 150  $\mu$ M malachite green (MG). Compound concentrations tested were 3.12 to 400  $\mu$ M (triplicate), unless otherwise specified.

MTB RNAPs were tested in the same fashion, using a 2x MTB reaction buffer solution (40 mM Tris-HCl (pH = 8.0 at 37°C), 20 mM MgCl<sub>2</sub>, 300 mM potassium L-glutamate, 15% glycerol, 20 nM pMGA12, 1 mM NTPs) with 80 nM pMGA12 or pMGA4-Mtb (Figure IV-2) plasmid. The concentration of the 2x RNAP solution was 80 nM MTB RNAP and 240 nM SigA.

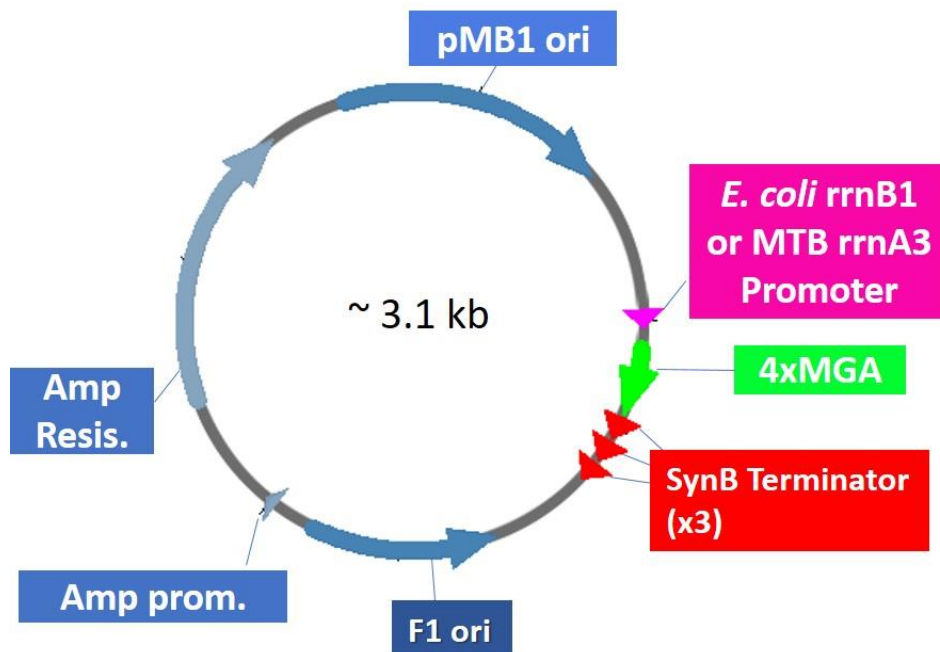


Figure IV-2. Plasmid map of DNA template containing terminator sequences and *E. coli* or MTB promoter. Plasmid containing *E. coli* rrnB1 promoter (pMGA4-Ec) is 3156 bp. Plasmid containing MTB rrnA3 promoter (pMGA4-Mtb) is 3179 bp.

The purchased compounds and synthesized compounds were both tested in this plasmid-based transcription assay as described. Data were analyzed and IC<sub>50</sub> values were determined as described in Chapter II, normalizing fluorescence values to percent activity and fitting data by nonlinear regression.

#### Assay to Detect Irreversibility of RNAP Inhibition

Compounds were tested for reversible inhibition of RNAP according to the scheme in Figure IV-3. Compounds (200  $\mu$ M) or DMSO were incubated with *E. coli* WT RNAP (2  $\mu$ M) at 37°C for 30 minutes in 1x *E. coli* reaction buffer without NTPs (20  $\mu$ L). After incubation, 4  $\mu$ L of the mixture was diluted 100-fold with 1x *E. coli* reaction buffer, diluting compounds 100-fold to 2  $\mu$ M, and the enzyme to 20 nM. Aliquots of 47  $\mu$ L were added to a 96-well plate, followed by the addition of 2  $\mu$ L compound or DMSO (in triplicate). Reactions were initiated by adding 1  $\mu$ L of 5

mM NTPs. After incubation at 37°C for 1.5 hours, activity was detected by fluorescence (Ex/Em = 628/660nm) after addition of 50  $\mu$ L of 150  $\mu$ M MG.

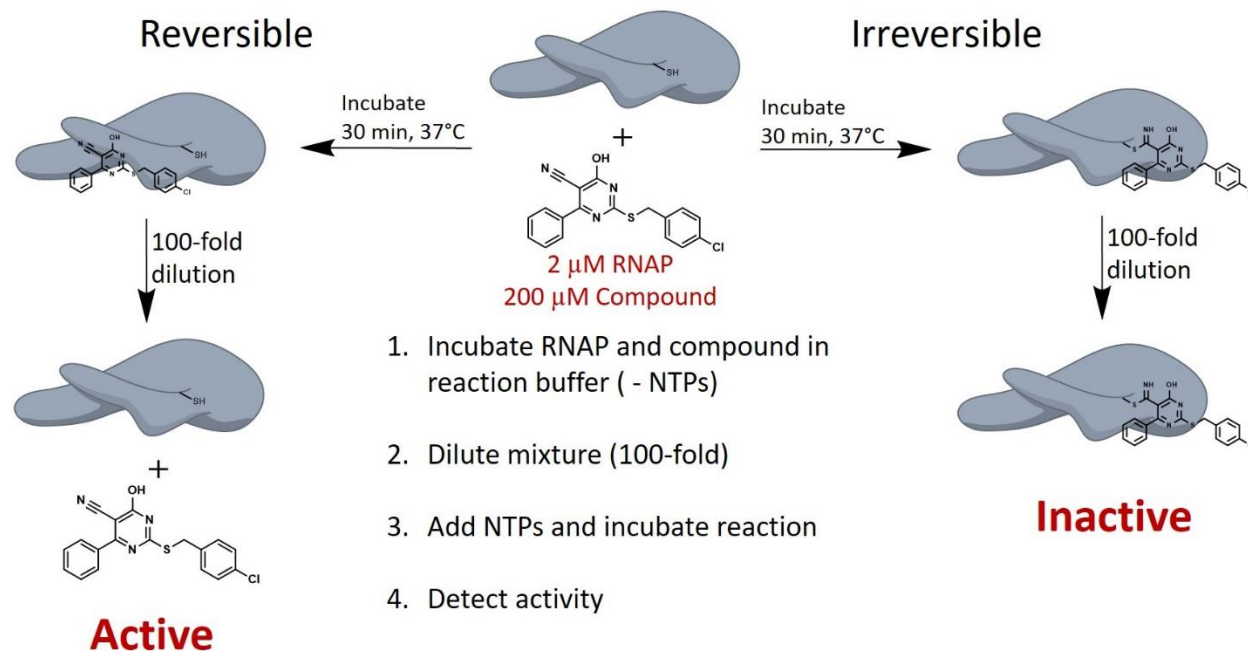


Figure IV-3. Overview of transcription assay used to detect reversibility of inhibition.

### Nitrile Reactivity Assay

The four cyanopyrimidine compounds (cluster 1 – 257176, 257180, 257181, 257182) identified from the high-throughput screen were tested for reactivity with L-cysteine as well as two control compounds, CCG-257084 and 5-bromopyrimidine-2-carbonitrile. CCG-257084 and 5-bromopyrimidine-2-carbonitrile have been shown to either not form or extensively form a cysteine adduct, respectively.<sup>100</sup> Compounds (or DMSO for background control) were incubated with L-cysteine to determine their ability to form adducts with L-cysteine, similarly to previously described for other aromatic nitrile compounds.<sup>100</sup> In short, reaction mixtures of 500  $\mu$ L were prepared containing 200  $\mu$ M compound, 2 mM L-cysteine, and 50 mM  $\text{Na}_2\text{HPO}_4$  (pH = 7.4), then incubated at 37°C for 30 min. The reaction mixtures were diluted in 1 mL of methanol, followed

by injection of 30  $\mu\text{L}$  onto an LC-MS instrument (Shimadzu) pre-equilibrated with 10:90 acetonitrile:water (v/v). The injected sample was then eluted over a 15-minute gradient from 10% to 90% acetonitrile (v/v). Pure compounds diluted in 500  $\mu\text{L}$  water and 1 mL methanol were also injected and analyzed as negative controls to help identify the original compound peak and potential contaminants or background signals.

## Results

### Fluorescent Intercalator Displacement Assay

To determine if compound inhibition of bacterial RNAP was potentially occurring via non-specific inhibition by DNA binding, the 15 active compounds identified from the HTS were evaluated for displacement of ethidium bromide from a DNA template. The compounds were tested at four different concentrations (12.5 – 100  $\mu\text{M}$ ) using the FID assay to determine if the compounds potentially inhibit RNA polymerase by binding DNA. Berenil (diminazine acetate),

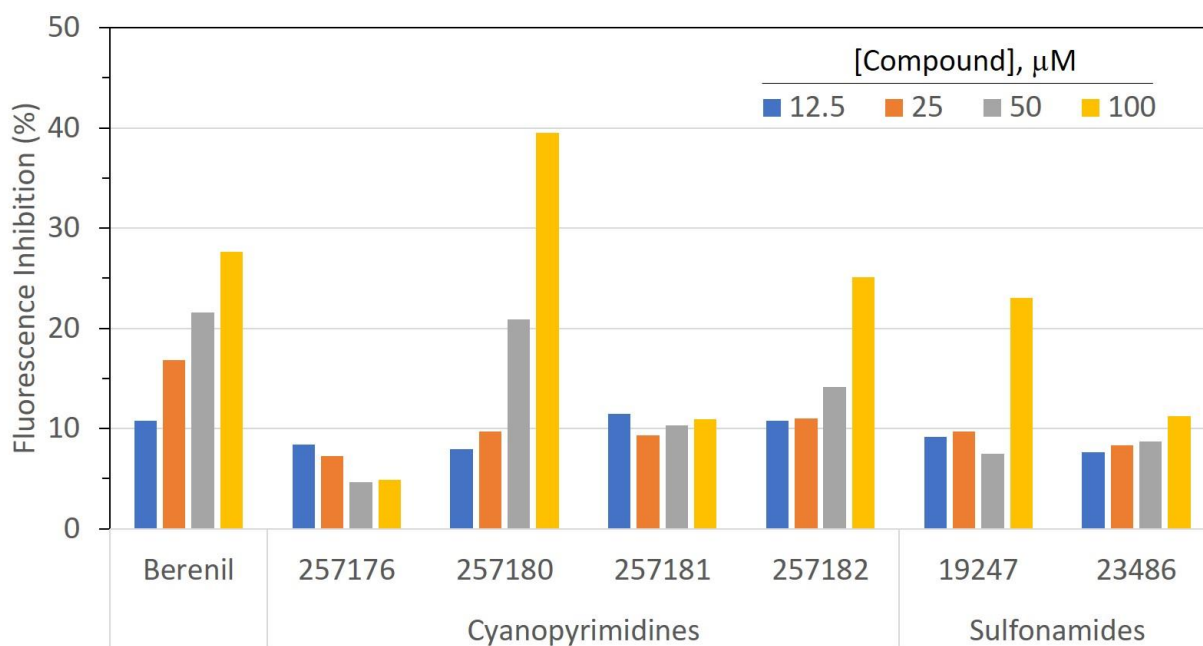


Figure IV-4. Fluorescence inhibition of ethidium bromide by compounds that were active in the fluorescent intercalator displacement (FID) assay. For comparison, data for FID inactive compounds (257176, 257181, and 23486) within the same cluster as the FID actives are also shown.

a known minor groove binder, was used as a control for assay validation.<sup>108</sup> Displacement of ethidium bromide by compounds was measured after incubation with a 60-bp double-stranded DNA probe (FID-lacUV5). In addition to the positive control (berenil), three of the compounds inhibited fluorescence of ethidium bromide in the FID assay at 100  $\mu$ M (one had some inhibition at 50  $\mu$ M also) using the FID-lacUV5 probe (Figure IV-4). However, these compounds also inhibited fluorescence of ethidium bromide when using a scrambled DNA sequence (data not shown), indicating that the binding is non-specific. For this and reasons in the Discussion section below, even though these compounds displace ethidium bromide, it seems likely that their inhibition of RNAP is not caused by binding to the DNA.

#### **MICs against virulent *M. tuberculosis* and Cytotoxicity**

The 15 compounds identified from the HTS were tested for inhibition of replicating (MABA) and non-replicating (LORA) *M. tuberculosis*. Of the compounds tested, only four compounds exhibited detectable growth inhibition at concentrations lower than 200  $\mu$ M (Table IV-2 and Figure IV-5), with toxicity IC<sub>50</sub> values of 50 – 100  $\mu$ M against Vero cells in the MTS assay.

Table IV-2. MICs of compounds active vs *M. tuberculosis* H<sub>37</sub>Rv

Compound	Cluster	MTB WT IC <sub>50</sub> ( $\mu$ M)	MIC <sub>90</sub> ( $\mu$ M)	
			MABA	LORA
257181	1	21	94	90
22785	5	53	85	81
186618	6	28	79	197
193157	6	>200	88	183
Rifampin		0.01	0.13	0.46

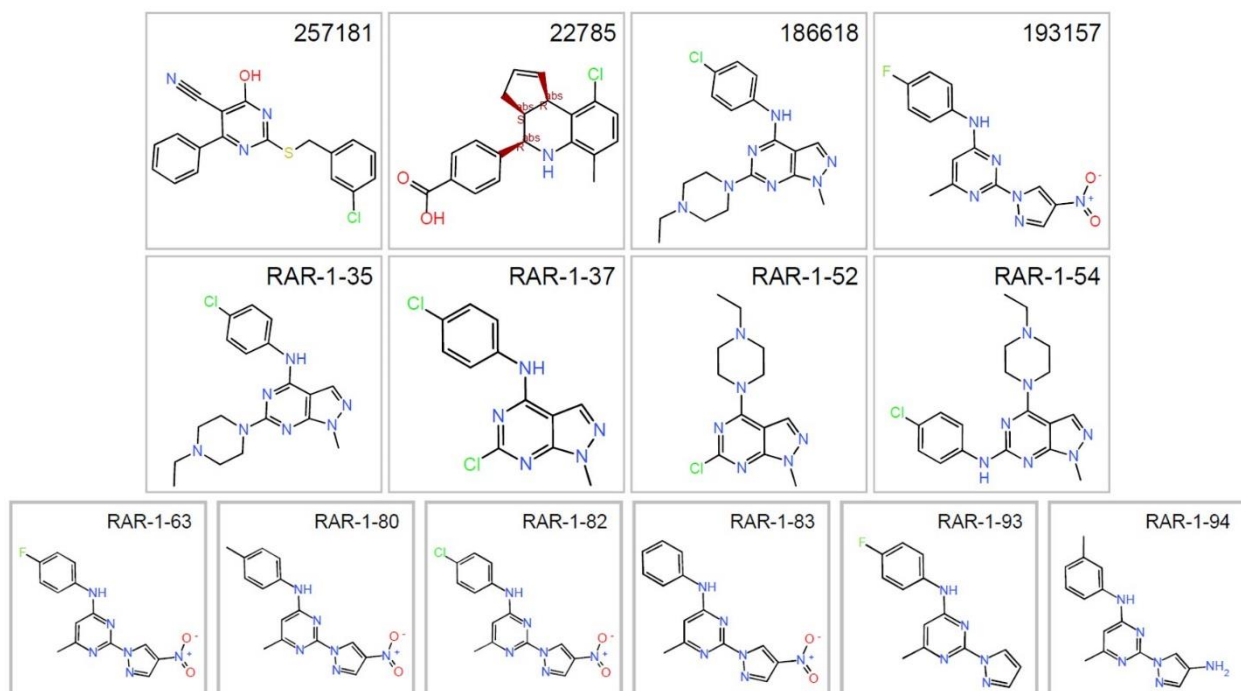


Figure IV-5. Structures of HTS hits active against MTB in culture and pyrazolopyrimidine analogs synthesized and tested for inhibition of *E. coli* and MTB RNAP. Top: HTS hits with activity against *M. tuberculosis* in LORA and/or MABA. Middle: Regioisomer and intermediates of HTS hit 186618. Bottom: Analogs of HTS hit 193157.

### Inhibition of RNAP by Pyrazolopyrimidines

A set of 30 commercially available analogs of the pyrazolopyrimidines (cluster 6, Appendix IV-1, Figure IV-A1) were purchased from ChemDiv (90% purity), and were tested for concentration-dependent inhibition against both *E. coli* and MTB WT RNAP at concentrations up to 200 and 100  $\mu\text{M}$  (respectively) using the plasmid-based transcription assay with pMGA12 as the plasmid template. Three compounds (PP-6, PP-16, and PP-22) had  $\text{IC}_{50}$  values less than 200  $\mu\text{M}$  against *E. coli* RNAP; however, the most potent compound against MTB RNAP only showed 33% inhibition at 100  $\mu\text{M}$  (PP-22).

The original hit compound 186618 and several analogs were synthesized and tested for concentration-dependent inhibition of *E. coli* and MTB RNAP, but were inactive against both enzymes (data not shown). The compound RAR-1-35 (re-synthesis of 186618) was inactive



against both *E. coli* and MTB RNAP. A potential alternative product (a structural isomer) was also prepared and tested for inhibition (RAR-1-54, Figure IV-5, top). Only a crude product of RAR-1-54 was obtained, but the crude product was found to be inactive against both *E. coli* and MTB RNAP. NMR and MS studies of the purchased compound (186618) and synthesized version (RAR-1-35) agree with one another, suggesting that the correct compound is the primary organic compound in solution. Analogs of 193157 were synthesized and tested for concentration-dependent inhibition of *E. coli* and MTB RNAP (Figure IV-5, bottom and Table IV-3).

Table IV-3. IC<sub>50</sub> values (μM) of 193157 analogs against *E. coli* WT RNAP

Compound	<i>E. coli</i>	MTB
RAR-1-63	54	>400
RAR-1-80	8.2	>400
RAR-1-82	12	>400
RAR-1-83	310	>400
RAR-1-93	320	>400
RAR-1-94	>400	>400

### Inhibition of RNAP by Sulfonamides

A set of analogs of 257178 and 23486 (11 and 10 analogs, respectively) were purchased from VitasM and ChemDiv (Figure IV-6 and Appendix IV-1, Figure IV-A2). The compounds were tested for concentration-dependent inhibition against both *E. coli* and MTB WT RNAP at concentrations up to 400 μM using the plasmid-based transcription assay with pMGA12 as the plasmid template. Only one analog of 23486 had an IC<sub>50</sub> value ≤ 400 μM (SA-20, IC<sub>50</sub> = 400 μM against MTB RNAP). Seven analogs of 257178 had MTB WT RNAP IC<sub>50</sub> values < 400 μM, ranging from 250 μM to 370 μM (Figure IV-6). None of the analogs of 257178 had IC<sub>50</sub> values < 400 μM against *E. coli* WT RNAP.

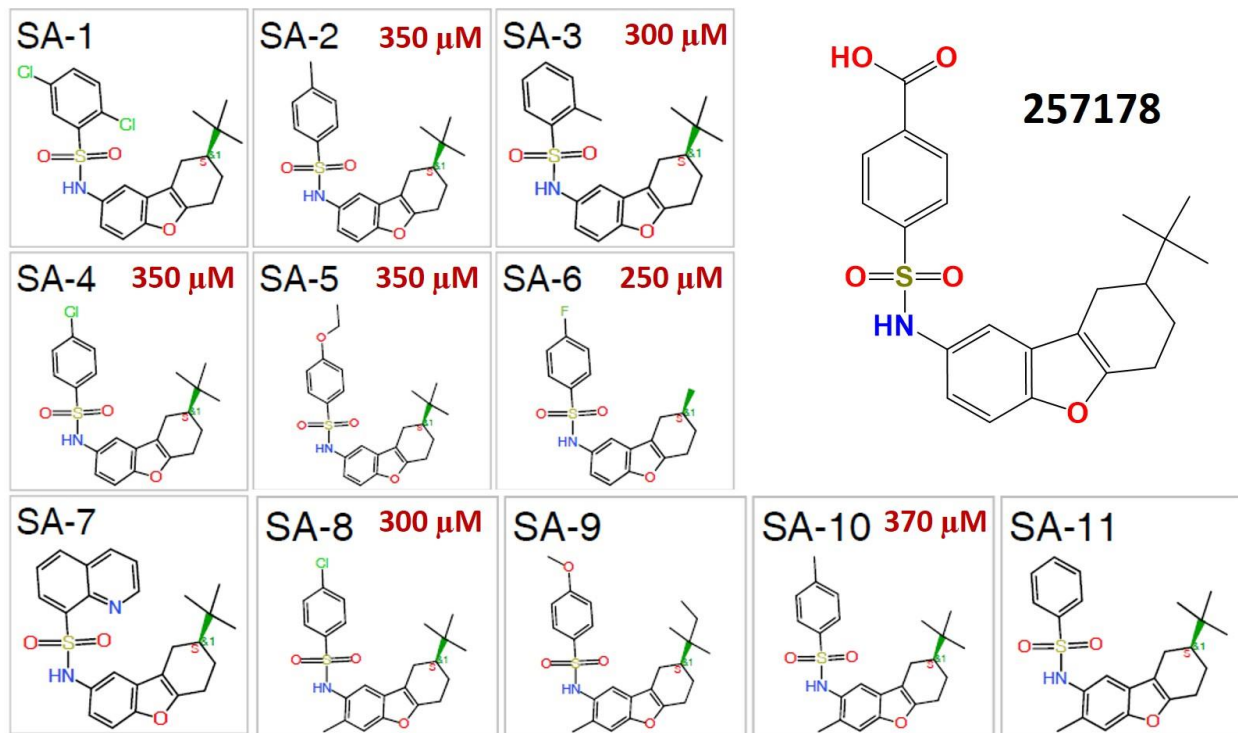


Figure IV-6. Analogs of HTS hit 257178 tested for inhibition against *E. coli* and MTB WT RNAPs in the plasmid-based transcription assay.  $IC_{50}$  values listed are against MTB WT RNAP, other compounds did not have a determinable  $IC_{50}$  against MTB WT RNAP. Compounds did not display 50% inhibition against *E. coli* WT RNAP at concentrations up to 400  $\mu$ M.

### Cyanopyrimidine Activity without Nitrile

Compounds with and without the aromatic nitrile at position  $R_4$  were purchased and tested for concentration-dependent inhibition of *E. coli* WT RNAP (Figure IV-7, Table IV-4). The compounds that lacked an aromatic nitrile did not inhibit RNAP at concentrations up to 200  $\mu$ M, while analogous compounds with the nitrile had determinable  $IC_{50}$  values less than 200  $\mu$ M (Table IV-4).

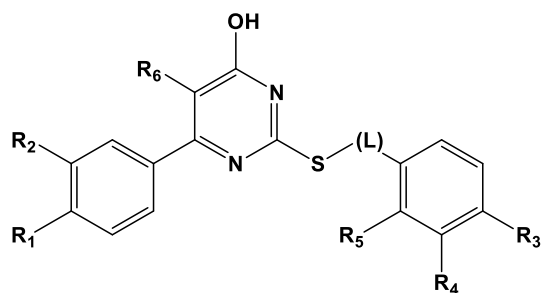


Figure IV-7. Structure of cyanopyrimidine scaffold tested for activity against *E. coli* and MTB WT RNAPs

Table IV-4. Activity of cyanopyrimidines +/- nitrile in plasmid-based transcription assay

Linker (L) = CH <sub>2</sub> R <sub>2</sub> = H, R <sub>5</sub> = H			IC <sub>50</sub> , μM	
R <sub>1</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>6</sub> = CN	R <sub>6</sub> = H
H	H	H	200	>200
H	H	Cl	50 (CN-1)	>200 (CN-2)
H	Cl	H	90	>200
OCH <sub>3</sub>	CN	H	150	ND
OCH <sub>3</sub>	H	H	100 (CN-3)	ND
H	CN	H	ND	>200

ND = Not determined

### Reversibility of Cyanopyrimidines

A couple representative compounds (CN-1, CN-2, CN-3 – Table IV-4) were tested for irreversible inhibition of *E. coli* WT RNAP by preincubating the compounds with RNAP, followed by a 100-fold dilution and transcription initiation. It was determined that RNAP retained transcriptional activity after the 100-fold dilution (Figure IV-8), indicating that the cyanopyrimidines are reversible inhibitors of bacterial RNAP.

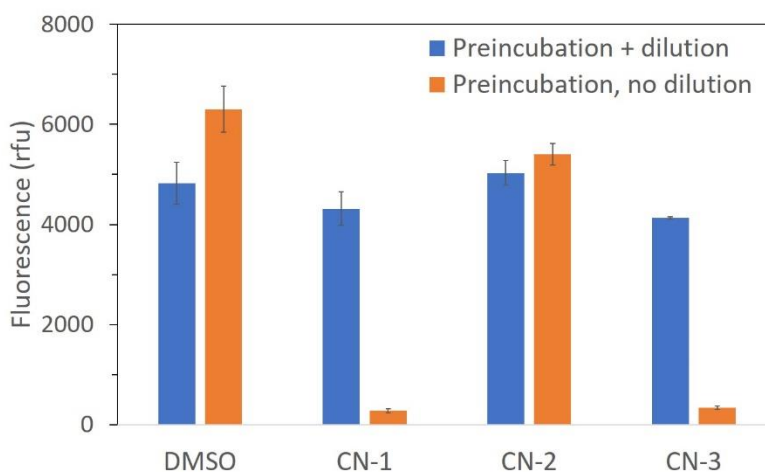


Figure IV-8. Plasmid-based transcription assay after RNAP preincubation with compounds. RNAP-compound mixtures were diluted 100-fold (from 200 μM compound to 2 μM, decreasing [compound] >10-fold below IC<sub>50</sub> values), followed by transcription initiation.

## Nitrile Reactivity Assay

The four cyanopyrimidines identified from the HTS were tested for reactivity with L-cysteine. None of the cyanopyrimidine compounds showed adduct formation, while the positive control 5-bromopyrimidine-2-carbonitrile was 100% converted to an adduct (Figure IV-9 and Appendix IV-3) with an  $m/z$  of 242 in negative ion mode, indicating an exact mass of 243. The 5-bromopyrimidine-2-carbonitrile/L-cysteine adduct had two peaks of similar intensity with  $m/z$  values of 242 and 244, indicating the presence of the two bromine isotopes, as expected.

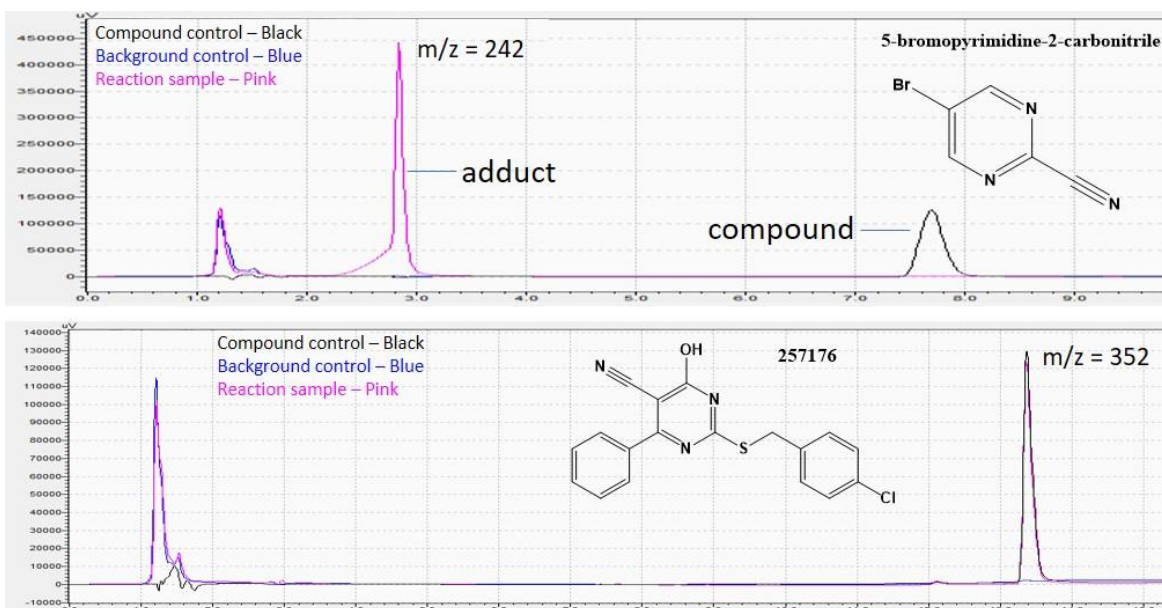


Figure IV-9. LC-MS chromatogram of positive control (5-bromopyrimidine-2-carbonitrile) and exemplary cyanopyrimidine (257176). The  $m/z$  values are from MS detection in negative ion mode.

## Inhibition of RNAP by Cyanopyrimidines

The compounds 257176 and 257181 identified from the HTS were synthesized and their inhibition of *E. coli* and MTB WT RNAP was confirmed in the plasmid-based transcription assay with pMGA12. A set of 27 cyanopyrimidine analogs (Figure IV-A3) were purchased and tested for inhibition against *E. coli* and MTB WT RNAPs using the plasmid-based transcription assay with pMGA4-Ec or pMGA4-Mtb, respectively. The compounds for this cyanopyrimidine scaffold

(Figure IV-7) were evaluated for concentration-dependent inhibition up to 400  $\mu\text{M}$  and  $\text{IC}_{50}$  values were determined for the compounds tested (Table IV-5). The compounds had  $\text{IC}_{50}$  values ranging from 5  $\mu\text{M}$  to > 400  $\mu\text{M}$ , indicating the potential for further SAR studies.

### Discussion

The top scaffolds of interest identified from a 150,000-compound high-throughput screen have been further studied and evaluated for their potential as bacterial RNAP inhibitors. Initial control screens during the HTS triage eliminated scaffolds that inhibited MG fluorescence or were not concentration-dependent. We further evaluated the scaffolds identified to validate their inhibition of bacterial RNAP and determine the potential of these scaffolds for future SAR studies. Research studying inhibition of DNA binding proteins by small molecules has utilized the FID assay to determine if the compounds potentially inhibit the protein by binding to DNA.<sup>105</sup> We utilized this approach and screened the 15 compounds identified for displacement of the fluorescent intercalator ethidium bromide. Only three of the compounds (257180, 257182, and 19247) inhibited fluorescence in the FID assay, which is indicative of DNA binding. However, this only occurred at concentrations more than 10-fold higher than their RNAP  $\text{IC}_{50}$ s. Therefore, it is reasonable to deduce that these compounds do not exhibit their RNAP inhibitory activities via DNA binding.

Nine compounds identified from the HTS had similar  $\text{IC}_{50}$  values across the panel of RNAP enzymes tested (Chapter III, Table III-4); however, only two of these compounds were active against *M. tuberculosis* in culture (257181 and 22785, Table IV-2). Two compounds from the pyrazolopyrimidine scaffold were also active in culture (186618 and 193157). Interestingly, compound 193157 was active in culture despite being inactive against MTB WT RNAP in vitro.

Table IV-5. Structure-activity relationship table of cyanopyrimidines against *E. coli* and MTB WT RNAPs

Compound	Linker	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	IC <sub>50</sub> (μM)	
		(para)	(meta)	(para)	(meta)	(ortho)	<i>E. coli</i>	MTB
257176	CH <sub>2</sub>	H	H	Cl	H	H	84	130
257180	CH <sub>2</sub>	H	H	H	Cl	H	55	80
CN-14	CH <sub>2</sub>	H	H	H	H	Cl	33	66
CN-12	CH <sub>2</sub>	H	H	H	H	2,6-Cl <sup>a</sup>	28	38
CN-25	CH <sub>2</sub>	H	H	F	H	H	47	110
CN-23	CH <sub>2</sub>	H	H	Br	H	H	31	37
CN-38	CH <sub>2</sub>	C(CH <sub>3</sub> ) <sub>3</sub>	H	Br	H	H	12	4.8
CN-18	CH <sub>2</sub>	H	H	H	CONH <sub>2</sub>	H	>400	>400
CN-34	CH <sub>2</sub>	Cl	H	COOH	H	H	>400	>400
CN-16	CH <sub>2</sub>	OCH <sub>3</sub>	H	Phenyl	H	H	33	15
CN-15	CH <sub>2</sub>	OCH <sub>3</sub>	H	H	H	Cl	28	45
CN-20	CH <sub>2</sub>	OCH <sub>3</sub>	H	Cl	H	Cl	24	27
CN-30	CH <sub>2</sub>	CH <sub>3</sub>	H	H	H	H	40	75
CN-27	CH <sub>2</sub>	CH <sub>3</sub>	H	CH <sub>3</sub>	H	H	28	38
CN-22	CH <sub>2</sub> CH <sub>2</sub> O	H	H	H	H	H	50	100
257181	CH <sub>2</sub> CH <sub>2</sub> O	Cl	H	OCH <sub>3</sub>	H	H	37	51
CN-36	CH <sub>2</sub> CH <sub>2</sub> O	Cl	H	Cl	H	H	18	15
CN-28	CH <sub>2</sub> CH <sub>2</sub> O	H	H	CH <sub>3</sub>	H	H	26	58
CN-33	CH <sub>2</sub> CH <sub>2</sub> O	H	H	OCH <sub>3</sub>	H	H	50	95
CN-31	CH <sub>2</sub> CH <sub>2</sub> O	CH <sub>3</sub>	H	CH <sub>3</sub>	H	H	22	31
CN-26	CH <sub>2</sub> CH <sub>2</sub> O	CH <sub>3</sub>	H	OCH <sub>3</sub>	H	H	28	50
CN-35	CH <sub>2</sub> CH <sub>2</sub> O	Cl	H	H	CH <sub>3</sub>	H	23	24
CN-37	(CH <sub>2</sub> ) <sub>3</sub> O	Cl	H	Cl	H	H	17	15
CN-24	(CH <sub>2</sub> ) <sub>4</sub> O	H	H	CH <sub>3</sub>	H	H	18	20
257182	(CH <sub>2</sub> ) <sub>4</sub> O	H	H	Cl	H	H	36	37
CN-17	CH <sub>2</sub> CC	OCH <sub>3</sub>	H	Cl	H	H	23	18
CN-13	CH <sub>2</sub> CO	CH <sub>3</sub>	H	Cl	H	H	160	94
CN-32	CH <sub>2</sub> COOCH <sub>2</sub>	Cl	H	H	H	H	>400	>400
CN-21	CH <sub>2</sub> COOCH <sub>2</sub> CH <sub>2</sub> O	H	H	H	H	H	160	290
CN-29	CH <sub>2</sub> CONHCH <sub>2</sub>	Cl	H	H	H	H	350	380

Structures are based on Figure IV-7 where R<sub>6</sub> = CN

<sup>a</sup>2,6-dichloro with Cl at both ortho positions on phenyl ring attached to linker (L)

It is possible that the activity of 193157 against *M. tuberculosis* is due to off-target or non-specific activity. However, it is known that *M. tuberculosis* does metabolize a number of drugs to more active forms<sup>109–112</sup>, therefore our compound may be metabolized (e.g., reduction of the nitro) to an active, RNAP-inhibiting derivative. Compound 193157 was synthesized and its activity against *E. coli* WT RNAP was confirmed. To further evaluate 193157 and the potential cellular metabolism of it, a set of analogs to 193157 were synthesized (Figure IV-5, bottom). The methyl and chloro substituents on the para position were more active than the fluoro; however, both were still inactive against MTB RNAP (Table IV-3). To evaluate the possibility that reduction of the nitro yields an active species against MTB WT RNAP, compounds with an amine (RAR-1-94) or without the nitro group (RAR-1-93) were synthesized (Figure IV-5, bottom) and evaluated for inhibition of *E. coli* and MTB WT RNAPs. The compound with the amine was inactive, and the compound without the nitro was weakly active against *E. coli* WT RNAP (Table IV-3). Since the amine was inactive, and all the analogs tested had IC<sub>50</sub> values near or > 400 μM, analogs of 193157 were not pursued further.

Compound 186618 was active against all RNAP enzymes tested, and active against replicating and non-replicating *M. tuberculosis*. This compound was synthesized, the structure and purity were confirmed (Rachel Rowlands and Dr. Andy White), and found to exhibit very low activity similar to other cluster 6 analogs. The NMR spectra and mass matched with the purchased compound and what was expected for the structure. Unfortunately, there was only a small amount of the purchased 186618 and it was no longer available from ChemDiv or other vendors, so we were unable to perform further analyses to determine the potential active component. The synthetic intermediates RAR-1-37 and RAR-1-52 were characterized and tested

for activity against *E. coli* and MTB RNAPs, with RAR-1-37 having an IC<sub>50</sub> near 200 μM (Table IV-3). The regioisomer RAR-1-54 was synthesized and did not inhibit *E. coli* or MTB RNAP at concentrations up to 200 μM. Based on these results, we suspect that the activity of the purchased sample of 186618 was due to a contaminant that we were unable to identify.

A set of sulfonamide analogs with structural similarity to 257178 or 23486 were evaluated for inhibition of *E. coli* and MTB WT RNAPs. Analogs of 257178 were all weakly active with IC<sub>50</sub> values of 300 μM or greater. The diversity of the commercially available analogs of 257178 is quite limited, mainly consisting of variations on the phenyl ring attached to the sulfonyl group. Additionally, 257178 is a charged molecule in aqueous solutions at physiological pH, with the negative carboxyl group on the phenyl ring. It is possible that this carboxyl group is important for binding, while the analogs tested do not contain the carboxyl group. A couple analogs have oxygen atoms that could act as a hydrogen bond donor (SA-5, SA-9), but do not have the potential for the same negative electrostatic interactions as a carboxyl group. This compound (or any analogs) was not synthesized to reconfirm and evaluate the activity due to the synthetic difficulty of this molecule compared to other scaffolds of interest. All the analogs of 23486 tested had IC<sub>50</sub> values ≥ 400 μM (Figure IV-A2), indicating that the scaffold is not optimal for an SAR campaign.

Initially, the cyanopyrimidine scaffold was of low interest due to the reported potential of these compounds for pan-assay interference via covalent interaction with cysteines.<sup>100</sup> However, over 30 nitrile-containing compounds are in clinical use, with additional lead compounds in clinical development.<sup>113</sup> The largest class of nitrile-containing drugs are compounds with an aromatic core containing a nitrile substituent, as is the case for the cyanopyrimidine scaffold identified from the HTS. These nitriles often act as hydrogen bond



acceptors with their therapeutic target; however, only a few have been shown to act covalently.<sup>114</sup> Given that numerous drugs containing aromatic nitriles are in clinical use, the mere presence of the aromatic nitrile should not block further development. Therefore, these nitrile containing compounds were subjected to additional investigation.

Testing a series of analogs that contained or lacked the aromatic nitrile (Table IV-4) revealed that analogs without the nitrile were essentially inactive. This result agreed with our original hypothesis, that the cyanopyrimidine compounds could potentially be inhibiting bacterial RNAP non-specifically by covalently interacting with cysteine. However, it is also possible that removing a potential hydrogen bond acceptor (nitrile) would decrease the activity of an inhibitor, so this result by itself did not determine if the cyanopyrimidines act via covalent inhibition. In order to confirm or refute this hypothesis, further testing was still necessary.

First, it was determined whether the compounds irreversibly inhibited *E. coli* WT RNAP by incubating the compounds with RNAP followed by dilution to reduce the concentration of compound significantly below the  $IC_{50}$ . Therefore, if the compound reversibly inhibits RNAP, the compound would be able to dissociate from RNAP and transcription would occur. As seen in Figure IV-8, RNAP that was preincubated with the active compounds CN-1 and CN-3 ( $IC_{50}$ s = 50 and 100  $\mu$ M, Table IV-4), then diluted 100-fold retained activity. The inactive compound CN-2 (lacking aromatic nitrile) was tested as a control, showing that the RNAP is active in the presence of CN-2 and that preincubation and dilution did not affect the enzyme. The results from these experiments show that the cyanopyrimidine compounds reversibly inhibit bacterial (*E. coli*) RNAP. However, these results do not refute the possibility that the aromatic nitrile might covalently interact with an off-target enzymatic cysteine resulting in toxicity.

The cyanopyrimidine compounds identified from the HTS (257176, 257180, 257181, and 257182) were tested for adduct formation with L-cysteine. This assay was utilized to determine the potential of the compounds to non-specifically, covalently interact with cysteine. It has been shown that depending on the electrophilicity of the aromatic nitrile, nitriles similar to the cyanopyrimidines can form cysteine adducts in solution with L-cysteine.<sup>100</sup> As shown in Figure IV-9 and Appendix IV-3, the compounds tested did not form cysteine adducts, indicating that the compounds are not electrophilic enough to interact with cysteines. The positive control used for this assay, 5-bromopyrimidine-2-carbonitrile, formed an adduct with an  $m/z$  of 242 in negative ion mode, indicating an exact mass of 243. The expected exact mass of the adduct that would form based on the proposed reaction in Figure IV-10 is 287. It is possible that the carboxylate from the cysteine was lost in the form of  $\text{CO}_2$  (MW = 44) to give an adduct with a mass of 243.

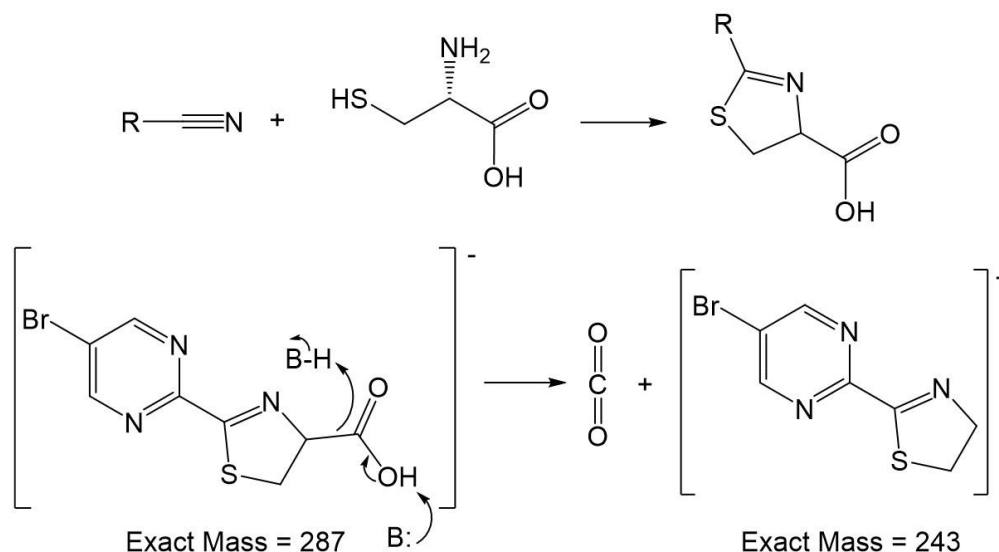


Figure IV-10. Proposed reaction for nitrile reactivity assay. Top: Reaction for generic nitrile with L-cysteine. Bottom: Proposed adduct formed and detected in electrospray ionization (negative mode).

Based on the data obtained for this cyanopyrimidine scaffold, these compounds are likely specific inhibitors of bacterial RNAP. The compounds inhibit both WT and  $\text{Rif}^{\text{R}}$  RNAPs for *E. coli* and MTB (Chapter III, Table III-4), therefore it is likely that they do not bind in the rifamycin

binding pocket. For future analogs of this scaffold, it is important to keep in mind the potential for them to be PAINS. If the core pyrimidine ring is changed where the electrophilicity of the nitrile may increase, such analogs should be tested in the nitrile reactivity assay. Iterative design, synthesis and testing is necessary to find more potent inhibitors of WT and Rif<sup>R</sup> MTB RNAP.

### Conclusions

The top scaffolds of interest identified from a high-throughput screen (Chapter III) were evaluated for their potential as a bacterial RNAP inhibitor. The pyrazolopyrimidine scaffold (Cluster 6) was eliminated because the most promising compound (186618) was not confirmed after synthesis, and an initial SAR effort with commercially available analogs did not identify additional active compounds. The sulfonamides (Clusters 2 and 3) were also eliminated because the initial SAR effort with commercially available analogs did not identify additional active compounds.

Based on initial experiments, the cyanopyrimidines (Cluster 1) require the nitrile group for activity; however, it was determined that they do not irreversibly inhibit *E. coli* RNAP in vitro. Furthermore, it was determined that the cyanopyrimidines do not form adducts with L-cysteine in the nitrile reactivity assay, further indicating that their mechanism of inhibition is not via non-specific interactions with a cysteine on bacterial RNAP. Based on these results and the initial SAR from commercially available analogs, this scaffold merits further study as a bacterial RNAP inhibitor.

## Notes to Chapter IV

### Acknowledgments

I would like to thank Dr. Andrew White for his expertise on synthetic chemistry, scaffold evaluation and valuable input for analog selection for scaffolds of interest. I would like to thank Rachel Rowlands for synthesis of compounds from the high-throughput screen for reconfirmation and scaffold evaluation studies (RAR compounds). I would like to thank Dr. Anthony Emanuele for his help and input on implementation of the FID assay. I would like to thank Joe Madak for his help with the optimization and analysis of the samples tested via LC-MS in the nitrile reactivity assay. I would like to thank Maxwell Stefan for construction of plasmid templates pMGA4-Ec and pMGA4-Mtb.

Portions of the work described in this chapter was published in *SLAS Discovery* in 2017 (Volume 22, Issue 3; pages 287 – 297).

### Abbreviations

DNA, deoxyribonucleic acid; FID, fluorescent intercalator displacement; HTS, high-throughput screen; IC<sub>50</sub>, half-maximal inhibitory concentration; LC-MS, liquid chromatography-mass spectrometry; LORA, low-oxygen recovery assay; MABA, microplate alamar blue assay; MG, malachite green; MIC, minimal inhibitory concentration; MS, mass spectrometry; MTB, *Mycobacterium tuberculosis*; MTS, 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium); NMR, nuclear magnetic resonance; NTP, nucleoside triphosphate; ori, origin of replication; PAIN, pan-assay interference compound; Rif<sup>R</sup>, rifampin-resistant ; RNA, ribonucleic acid; RNAP, RNA polymerase; SAR, structure-activity relationships; TB, tuberculosis; WT, wild-type

## Appendix IV-1

### Purchased Analogs of Scaffolds Identified from HTS

A set of commercially available analogs were purchased for each scaffold of interest identified from the HTS. A set of 30 analogs of 186618 (pyrazolopyrimidine, cluster 6, Figure IV-A1) were purchased from ChemDiv and tested against *E. coli* and MTB WT RNAPs in the plasmid-based transcription assay using pMGA12. The IC<sub>50</sub> values in Figure IV-A1 are against *E. coli* WT RNAP. None of the compounds reached 50% inhibition at the highest concentration tested against MTB WT RNAP (100  $\mu$ M).

A set of analogs of 257178 (Figure IV-6) and 23486 (Figure IV-A2) were purchased from ChemDiv (90% purity) and VitasM (92% purity). Each analog was tested against *E. coli* and MTB WT RNAPs in the plasmid-based transcription assay using pMGA12. The IC<sub>50</sub> values in Figure IV-A1 are against MTB WT RNAP. None of the compounds reached 50% inhibition at the highest concentration tested against *E. coli* WT RNAP (400  $\mu$ M).

A set of 27 analogs for the cyanopyrimidine scaffold (Figure IV-A3) were purchased from various suppliers and tested for inhibition of *E. coli* and MTB WT RNAPs in the plasmid-based transcription assay using pMGA4-Ec or pMGA4-Mtb, respectively. The IC<sub>50</sub> values for the compounds ranged from 5  $\mu$ M to > 400  $\mu$ M against *E. coli* and MTB WT RNAPs.

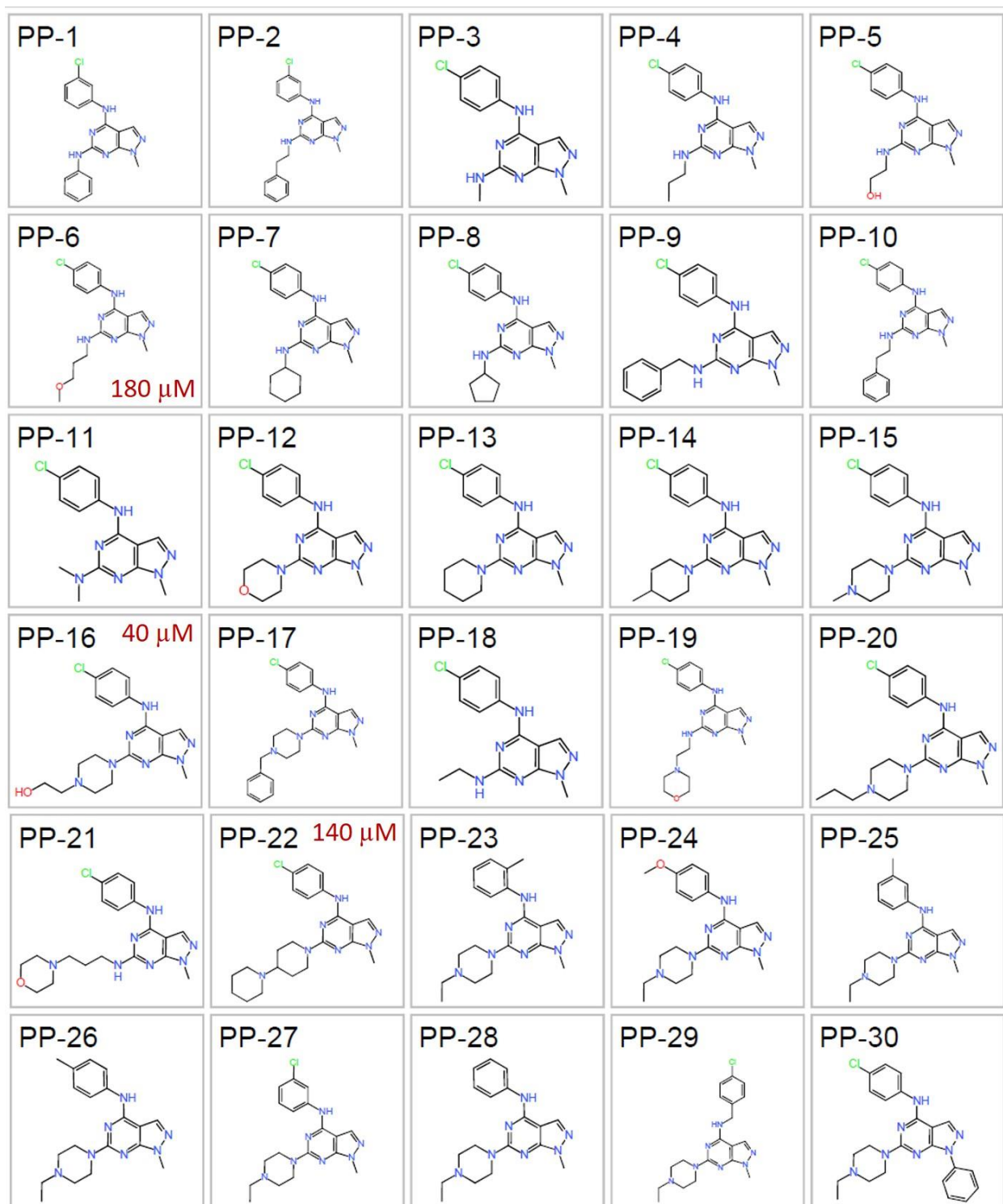


Figure IV-A1. Pyrazolopyrimidine analogs tested for inhibition against *E. coli* and MTB WT RNAPs in the plasmid-based transcription assay. IC<sub>50</sub> values listed are against *E. coli* WT RNAP, other compounds did not have a determinable IC<sub>50</sub> against *E. coli* WT RNAP. Compounds did not display 50% inhibition against MTB WT RNAP at concentrations up to 100  $\mu$ M.

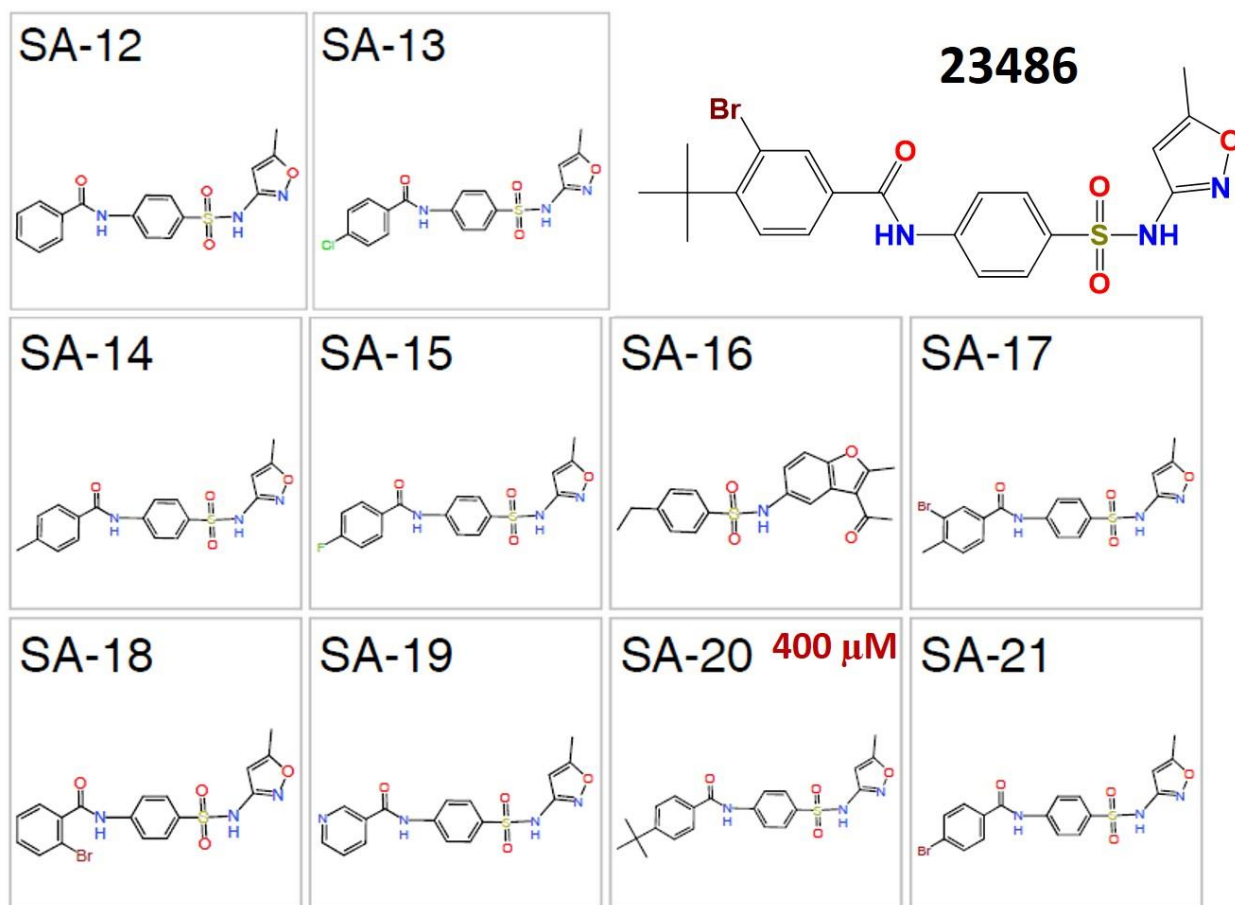


Figure IV-A2. Analogs of 23486 tested for inhibition against *E. coli* and MTB WT RNAPs in plasmid-based transcription assay. IC<sub>50</sub> value listed is against MTB WT RNAP, other compounds did not have a determinable IC<sub>50</sub> against MTB WT RNAP. Compounds did not display 50% inhibition against *E. coli* WT RNAP at concentrations up to 400 μM.

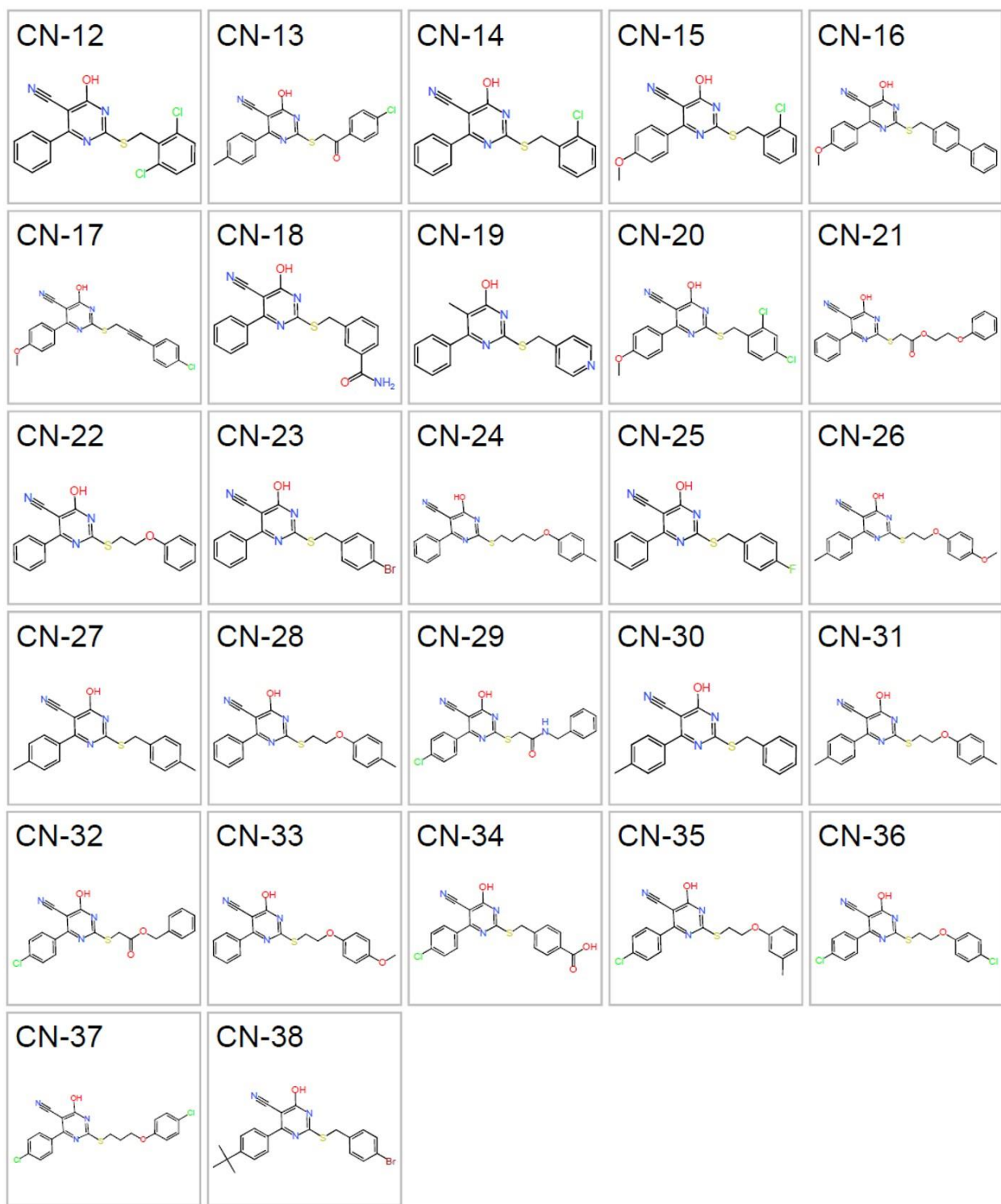


Figure IV-A3. Analogs of cyanopyrimidine scaffold tested for inhibition against *E. coli* and MTB WT RNAPs in plasmid-based transcription assay.



## Appendix IV-2

### Chemistry and Synthesis

Compounds were synthesized by Rachel Rowlands using the reaction schemes in Figures IV-A4 and IV-A5 for the 186618 and 193157 pyrazolopyrimidine analogs and intermediates, and Figure IV-A6 for the cyanopyrimidine compounds.

$^1\text{H}$  NMR spectra were recorded on Bruker 500 MHz spectrometers. Chemical shifts are reported in  $\delta$  (parts per million, ppm), by reference to the hydrogen residues of deuterated solvent as internal standard ( $\text{CDCl}_3$ :  $\delta = 7.28$  ppm;  $\text{DMSO-d}_6$ :  $\delta = 2.50$  ppm). Signals are described as s, d, t, q, m, dd, dt, ddt for singlet, doublet, triplet, quartet, multiplet, doublet of doublets, doublet of triplets, and doublet of doublet of triplets, respectively. Coupling constants (J) are given in Hertz (Hz).

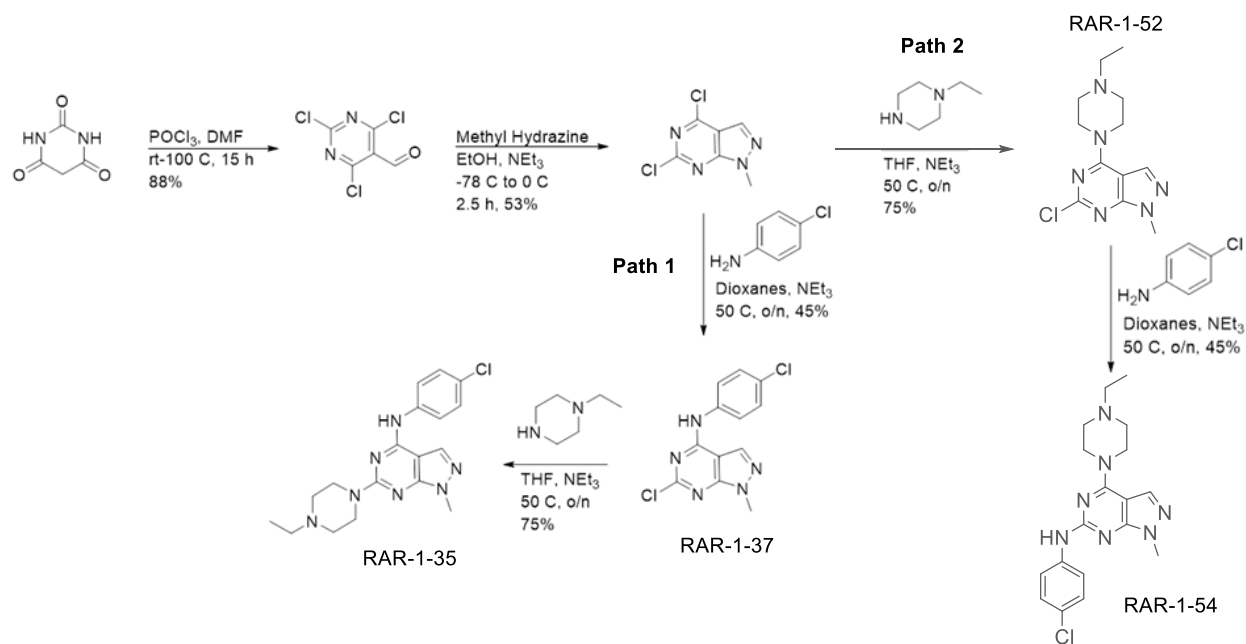


Figure IV-A4. Synthetic scheme for RAR-1-35, RAR-1-54, and intermediates.

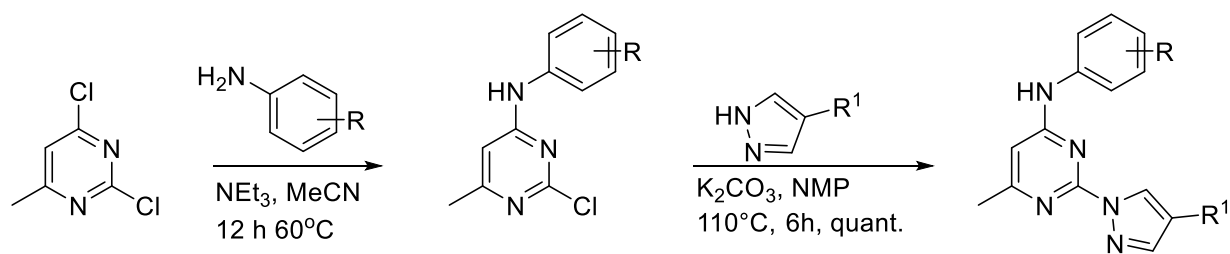


Figure IV-A5. Synthetic scheme for 193157 analogs and intermediates.

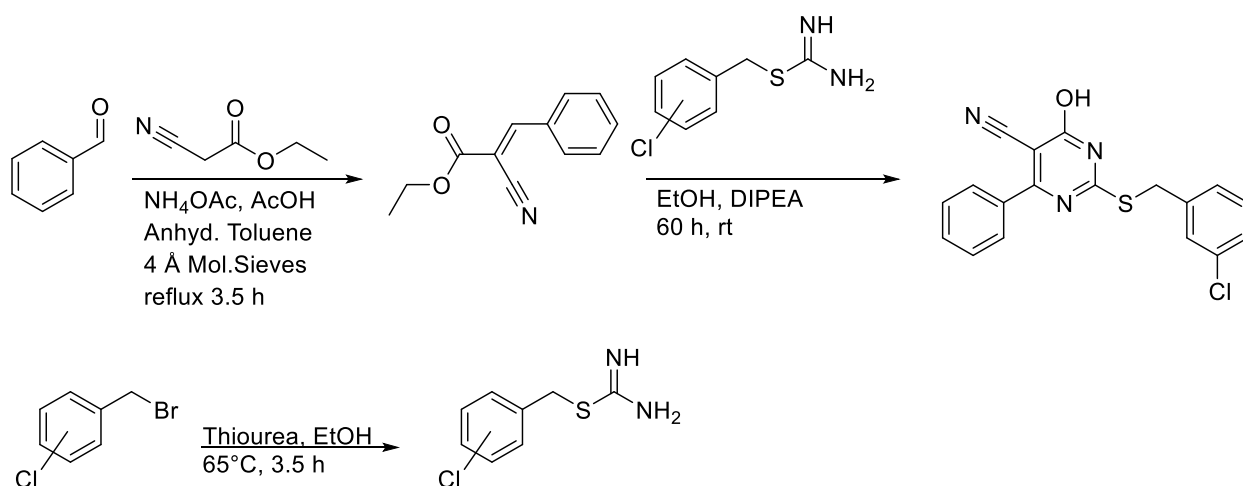


Figure IV-A6. Synthetic scheme for cyanopyrimidines 257176 and 257181.

#### RAR-1-35 (186618)

MS:  $m/z$  372.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.16 (s, 3H), 2.55 (s, 7H), 3.86 (s, 3H), 3.94 (s, 3H), 6.78 (s, 1H), 7.42 (d, 2H), 7.49 (s, 1H), 7.51 (d, 2H).

#### RAR-1-37

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.84 (s, 1H), 7.54 (d, 2H), 7.52 (s, 1H), 7.43 (d, 2H), 4.12 (s, 3H).

**RAR-1-52**

$^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ ):  $\delta$  8.4 (s, 1H), 3.8 (s, 3H), 3.6 (m, 2H), 3.1 (m, 4H), 1.3 (m, 4H), 1.2 (m, 3H).

**RAR-1-54**

MS:  $m/z$  352.  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ ):  $\delta$  8.38 (s, 1H), 7.06 (d, 3H), 6.96 (d, 4H), 3.92 (s, 3H), 1.23 (m, 11H), 1.08 (m, 7H), 0.88 (m, 10H).

**RAR-1-63 (193157)**

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.28 (s, 1H), 8.338 (s, 1H), 7.33 (t, 2H), 7.19 (t, 2H), 6.98 (s, 1H), 6.42 (s, 1H), 2.46 (s, 3H).  $^{19}\text{F}$  NMR (500 MHz,  $\text{CDCl}_3$ ): -120.

**RAR-1-80**

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.26 (s, 1H), 8.30 (s, 1H), 7.24 (d, 2H), 7.16 (d, 2H), 7.04 (s, 1H), 6.48 (s, 1H), 2.43 (s, 3H), 2.39 (s, 3H).

**RAR-1-82**

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.28 (s, 1H), 8.34 (s, 1H), 7.45 (d, 2H), 7.33 (d, 2H), 7.06 (s, 1H), 6.53 (s, 1H), 2.51 (s, 3H).

**RAR-1-83**

$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.32 (s, 1H), 8.31 (s, 1H), 7.85 (s, 1H), 7.50 (m, 2H), 7.28 (m, 3H), 6.62 (s, 1H), 2.46 (s, 3H)

**RAR-1-93**

MS:  $m/z$  270.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.58 (d, 1H), 7.82 (s, 1H), 7.32 (q, 2H), 7.16 (t, 2H), 6.94 (s, 1H), 6.48 (s, 1H), 6.34 (s, 1H), 2.45 (s, 3H).  $^{19}\text{F}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  -120.

**RAR-1-94**

MS:  $m/z$  281.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.10 (s, 1H), 7.64 (s, 1H), 7.50 (s, 1H), 7.32 (m, 1H), 7.16 (m, 1H), 7.10 (s, 1H), 7.04 (d, 1H), 6.45 (s, 1H), 2.43 (s, 6H, incidental Me overlap)

$^{13}\text{C}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  168.10, 162.43, 155.51, 129.54, 137.97, 136.00, 129.97, 125.98, 123.61, 119.99, 115.42, 98.98, 24.30, 21.42

**257176**

MS:  $m/z$  354.  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  7.93 – 7.88 (m, 2H), 7.63 (t,  $J = 7.3$  Hz, 1H), 7.58 (dd,  $J = 8.2, 6.4$  Hz, 2H), 7.51 (d,  $J = 1.9$  Hz, 1H), 7.39 (dt,  $J = 7.2, 1.8$  Hz, 1H), 7.36 – 7.30 (m, 2H), 4.52 (s, 2H).

**257181**

MS:  $m/z$  354.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.89 (dt,  $J = 12.5, 7.9$  Hz, 2H), 7.58 (ddt,  $J = 16.8, 12.5, 7.3$  Hz, 3H), 7.49 – 7.29 (m, 4H), 4.55 – 4.46 (m, 2H).

## Appendix IV-3

### Nitrile Reactivity Assay

The positive and negative controls 5-bromopyrimidine-2-carbonitrile and CCG-257084, respectively, were used for assessing nitrile reactivity for this assay based on data previously reported.<sup>100</sup> Four cyanopyrimidines identified from the HTS were tested for reactivity with L-cysteine in this aqueous assay. The reactions were applied to a pre-equilibrated Shimadzu LC-MS. The samples were separated and eluted using a 10% to 90% acetonitrile in water gradient over the course of 15 min. The chromatograms reported in Figure IV-A4 are the Absorbance intensity as a function of time in minutes. The UV absorption is measured using a photodiode array at a wavelength of 254 nm.

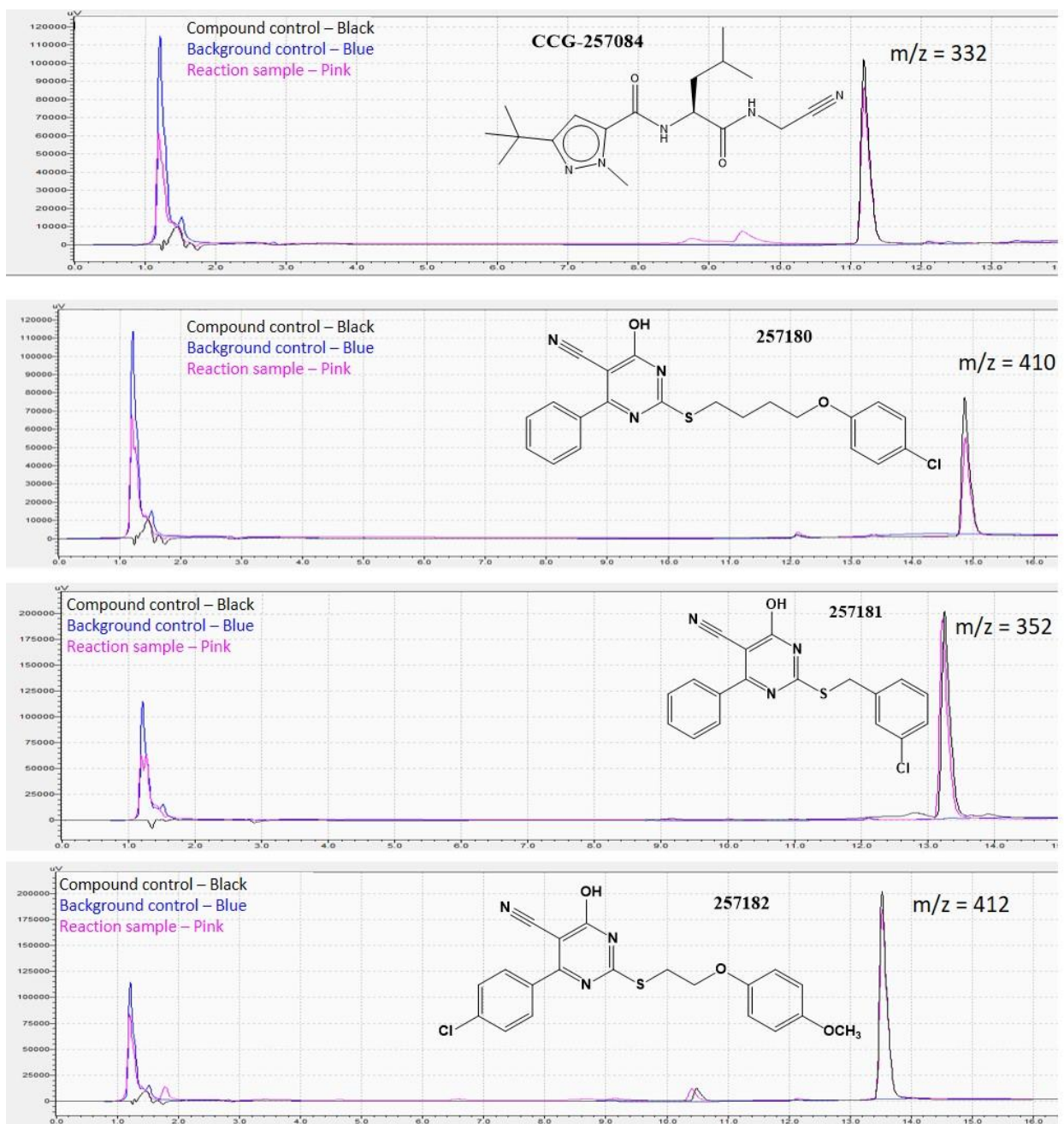


Figure IV-A7. LC-MS chromatogram of negative control CCG-257084 and cyanopyrimidines 257180, 257181, and 257182. The  $m/z$  values are from MS detection in negative ion mode.

## Chapter V

### **In vitro Evaluation of Benzoxazinorifamycin Analogs against Wild-type and Rifamycin-Resistant Mutants of *Escherichia coli* and *Mycobacterium tuberculosis* RNA Polymerase**

There have been few improvements in the treatment success and global burden of multidrug-resistant TB (MDR-TB) over the past decade. Each year, there are approximately half a million new cases reported.<sup>2</sup> Cases of MDR-TB are resistant to the two first-line drugs used throughout treatment, rifampin and isoniazid. Due to the resistance of *M. tuberculosis* to these drugs, MDR-TB treatment consists of second-line injectable anti-TB drugs such as fluoroquinolones and aminoglycosides.<sup>115,116</sup> In addition to the cost and issues related to intravenous administration of drugs, the length of therapy is 9 to 12 months (previously 18 to 24 months) instead of a 6-month treatment for drug-susceptible TB.<sup>2</sup> It is therefore of interest to identify agents that are able to overcome the issues of drug-resistant TB.

Rifampin-resistant (Rif<sup>R</sup>) TB is primarily caused by single amino acid mutations in a portion of the *rpoB* gene that encodes the  $\beta$  subunit of RNAP where rifampin binds. Rifamycins inhibit RNAP transcription initiation by preventing the formation of nascent RNA longer than 2 or 3 nucleotides, but they do not inhibit RNA transcription in the elongation phase.<sup>117</sup> The three most prevalent amino acid mutations are D516V, H526Y, and S531L (*E. coli* numbering). Rifampin and several other rifamycins have been tested in vitro against RNAP with these clinically relevant mutations, and it was shown that the Rif<sup>R</sup> mutants are approximately 10<sup>4</sup> to 10<sup>5</sup>-fold less susceptible to rifampin.<sup>56</sup> However, this work also demonstrated that the rifamycin analog

rifabutin was 10- to 50-fold more potent than rifampin against D516V and S531L *E. coli* RNAP. Therefore, it appears that additional improvements to the activity against Rif<sup>R</sup> mutants can be achieved through improved design of rifamycin analogs.

In addition to the issues of rifamycin resistance, rifampin is also a potent inducer of CYP3A4 by activating the human pregnane X receptor (hPXR).<sup>47</sup> Rifalazil, a benzoxazinorifamycin (bxRif) analog developed and tested in phase I and II clinical trials, was of great interest for several reasons. It is an exceedingly potent rifamycin with high tissue affinity, a longer half-life than rifampin, and had a relative lack of toxicity in early rodent studies.<sup>46,48,52</sup> Additionally, rifalazil did not show induction of CYPs or activation of hPXR as is seen for rifampin.<sup>49,57</sup> Unfortunately, rifalazil was shown to be toxic in phase I and phase II clinical trials with a flu-like syndrome and

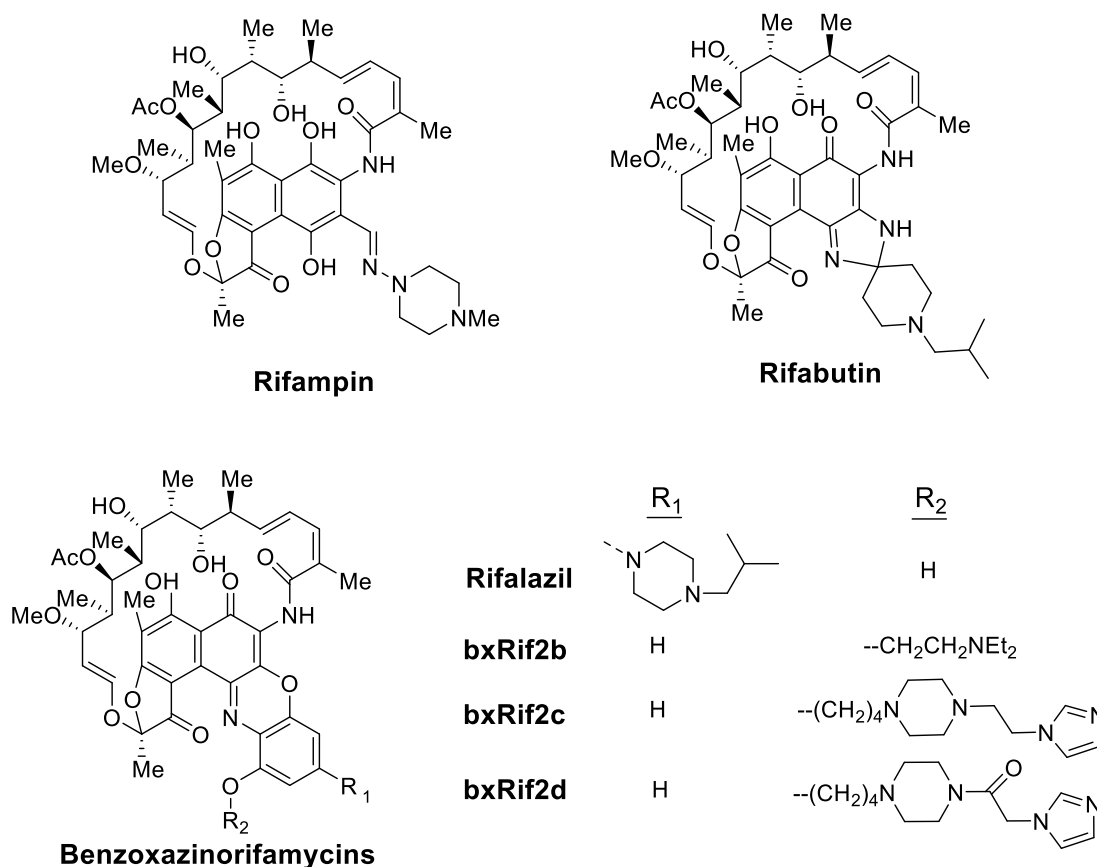


Figure V-1. Structures of benzoxazinorifamycins and the clinical rifamycins rifampin and rifabutin.



leucopenia, even at lower dosage levels.<sup>52,118,119</sup> Therefore, its development for the treatment of TB was suspended.<sup>44</sup> Aside from the toxicity issues of rifalazil, it demonstrated that a more potent rifamycin analog without the issues of drug-drug interactions can be developed.

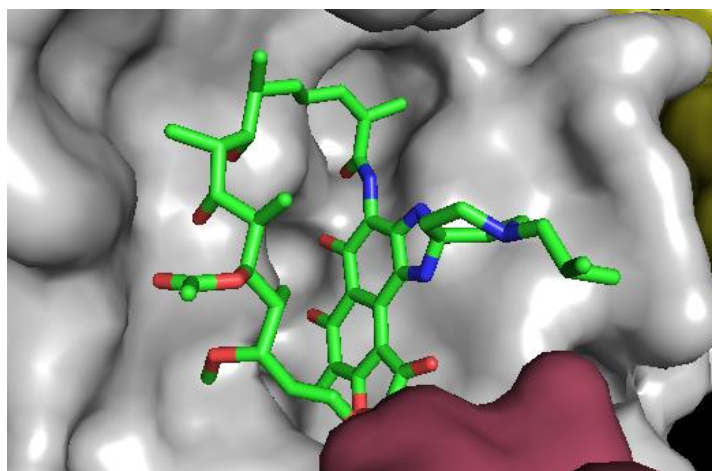


Figure V-2. Rifabutin bound to *Thermus thermophilus* RNAP (PDB: 2A68). RNAP molecular surface is shown with  $\beta$  subunit colored grey,  $\beta'$  in gold, and sigma factor in red.

Rifalazil and the bxRifs are a newer generation of rifamycin derivatives with a four-ring structure (Figure V-1). The bxRifs have shown improved antimicrobial activity against *M. tuberculosis* (MTB) and in vitro activity against Rif<sup>R</sup> RNAPs.<sup>48,57,120</sup> Prior work with the bxRifs resulted in the development of bxRif analogs 2b – 2d using modeling based on the 2.5 Å resolution structure of rifabutin (Figure V-1) in complex with the *Thermus thermophilus* RNAP holoenzyme (Figure V-2, PDB: 2A68).<sup>57</sup> These analogs were developed with the goal of interacting with the  $\sigma$  hairpin loop and other portions of the RNAP complex. The bxRif analogs prepared showed improved activity in vitro; however, the structure-based design utilized the *Thermus thermophilus* structure. The sequence and antibiotic sensitivities of the *E. coli* RNAP are closer to those of the MTB RNAP than are those of *T. thermophilus* RNAP, i.e., the *E. coli* RNAP is a better model to study RNAP – inhibitor interactions for TB drug discovery.<sup>117</sup> Recent work on bacterial RNAP has resulted in the structural determination of *E. coli* RNAP holoenzyme (PDB:

4YG2).<sup>95</sup> Therefore, the structure of rifampin and bxRif analogs in complex with *E. coli* WT RNAP holoenzyme was determined to evaluate their binding and interactions, and correlate the in vitro activities to the structural data.<sup>96</sup>

To further understand the binding of rifamycins to Rif<sup>R</sup> RNAP and the effects of these Rif<sup>R</sup> mutations (D516V, H526Y, S531L – *E. coli* numbering), we evaluated the rifampin inhibition of these mutants for *E. coli* and MTB RNAPs as well as recent X-ray crystal structures of *E. coli* Rif<sup>R</sup> RNAPs in complex with rifampin in collaboration with the Murakami lab (Penn State University). With the goal of identifying a potent rifamycin analog against MTB in culture as well as WT and Rif<sup>R</sup> RNAP, we developed and tested a series of novel bxRif analogs for their in vitro and cellular activity.

## **Materials and Methods**

### **Materials**

All the materials were the same as those specified in Chapter II. Oligonucleotides used in this study were purchased from Invitrogen (Table V-1). Quant-iT RiboGreen was purchased from ThermoFisher (Waltham, MA). Reagents used for rolling circle transcription assay were purchased from Epicentre (Madison, WI). The transcription plasmid templates were purified as described for pMGA12 in Chapter III. The crystallography work was done by Drs. Katsu Murakami and Vadim Molodtsov. Synthesis of bxRifs was done by Dr. Walajapet Rajeswaran based on methods previously published.<sup>57</sup> DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility.

Table V-1. List of oligonucleotides used

Oligonucleotide Name	Sequence, 5' - 3'	Length
pRSF-T7Del-Fwd	GCGACTCCTGCATTAGGCAGGTCGACAAGCTTGC	34
pRSF-T7Del-Rev	GCAAGCTTGTGACCTGCCTAATGCAGGAGTCGC	34
EcSig70-NdeI-Mut-Fwd	CCGTATTCCGGTGCACATGATTGAGACCATCAACAAG	37
EcSig70-NdeI-Mut-Rev	CTTGTTGATGGTCTCAATCATGTGCACCGGAATACGG	37

### Purification of *E. coli* and MTB RNAPs

For expression of the *E. coli* RNAP, a plasmid containing *E. coli*  $\sigma^{70}$  (without His tag) that was compatible with the pVS10 expression vectors (for co-expression) was prepared using the pRSFDuet-1 vector, pET15b-Sig70, and the oligonucleotides listed in Table V-1. One of the T7 promoters on the co-expression plasmid pRSFDuet-1 was removed via two-step site-directed mutagenesis using the complementary oligonucleotides pRSF-T7Del-Fwd and pRSF-T7Del-Rev to give pRSF-1. The plasmid was verified by DNA sequencing. An *NdeI* restriction site in the *E. coli* *rpoD* gene (Sig70) was mutated (silent mutation to preserve protein sequence) using site-directed mutagenesis with the complementary oligonucleotides EcSig70-NdeI-Mut-Fwd and EcSig70-NdeI-Mut-Rev. The resulting plasmid and pRSF-1 were digested with *NdeI* and *EcoRV*, then gel purified. The *E. coli* *rpoD* fragment and vector pRSF-1 were ligated to give the expression plasmid pRSF-Sig70. Plasmid was verified by DNA sequencing.

*E. coli* RNAP holoenzymes (WT, D516V, H526Y, S531L,  $\beta$ S531L/ $\beta$ 'V408G) were expressed and cells harvested as described in Chapter II, with cells containing expression vectors pVS10 and pRSF-Sig70 grown and expressed in terrific broth with 1 mM ZnCl<sub>2</sub> instead of 2xTY. MTB RNAP holoenzymes (WT, D435V, H445Y, S450L,  $\beta$ S450L/ $\beta$ 'V483G) were expressed and cells harvested as described in Chapter II, with cells grown and expressed in terrific broth with 1 mM ZnCl<sub>2</sub> instead of 2xTY.

*E. coli* RNAPs were purified as follows. Cell pellet from 1 L of culture was resuspended in 20 mL of RNAP lysis buffer (20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20  $\mu$ M ZnCl<sub>2</sub>, 5% glycerol, 2 mM  $\beta$ -mercaptoethanol, 1 mM PMSF and 1X Roche cOMplete ULTRA protease cocktail). Cells were lysed by sonication and clarified by centrifugation at 25,000 x g for 45 min. RNAP was precipitated by addition of 10% polymine P (pH 7.9) to a final concentration of 0.6% polymine P, followed by centrifugation at 6,000 x g for 10 min. The pellet was washed by resuspension in 12 mL RNAP lysis buffer containing 500 mM NaCl, followed by centrifugation (6,000 x g, 10 min, 4°C). RNAP was then solubilized with 20 mL RNAP lysis buffer containing 1 M NaCl, and clarified by centrifugation (8,000 x g, 10 min, 4°C). RNAP was precipitated from the supernatant by addition of 0.3 g ammonium sulfate per mL supernatant (6 grams), followed by incubation at 4°C for 30 min. After centrifugation (10,000 x g, 30 min, 4°C), RNAP was resuspended in ~5 mL dialysis buffer and dialyzed overnight (20 mM Tris-HCl (pH 8.0), 75 mM NaCl, 20  $\mu$ M ZnCl<sub>2</sub>, 5% glycerol, 2 mM  $\beta$ -mercaptoethanol). Dialyzed RNAP was sterile filtered and applied to a 1-mL HisTrap HP column (GE Healthcare) and purified by FPLC (ÄKTA, GE Healthcare). Isolated holoenzyme was then further purified on a 5-mL Source 15S column (cation exchange), followed by a 5-mL Source 15Q column (anion exchange).

MTB RNAPs were purified in the same manner as *E. coli* RNAPs with the following modifications. After ammonium sulfate precipitation and dialysis, MTB RNAP was applied to the 5-mL Source 15S column, followed by the 1-mL HisTrap column, then the 5-mL Source 15Q column. Purified MTB RNAP fractions were collected, dialyzed into storage buffer (40 mM Tris-HCl (pH 7.9), 200 mM NaCl, 0.1 mM EDTA, 1 mM MgCl<sub>2</sub>, 20  $\mu$ M ZnCl<sub>2</sub>, 50% glycerol, and 5 mM DTT), and stored at -20°C. *E. coli*  $\sigma^{70}$  and MTB SigA were purified as described in Chapter III.

### **Rolling Circle Transcription Assay (RiboGreen Detection)**

The activity of *E. coli* WT and Rif<sup>R</sup> RNAPs were measured via rolling circle transcription assay using the Kool NC-45 Universal RNA Polymerase template in the absence and presence of bxRifs. Concentration-response studies with rifalazil (bxRIF2a) and bxRif analogs (bxRif2b–d) were performed via rolling circle transcription assay in the same manner as previously described to determine IC<sub>50</sub> values against *E. coli* RNAPs. In brief, core *E. coli* RNAP (10 nM for WT, 100 nM for mutants) was incubated with the test compound for 10 min at 37 °C in a 25 µL reaction volume of 1× RNAP reaction buffer (40 mM Tris-HCl (pH 8.0), 50 mM KCl, 10 mM MgCl<sub>2</sub>, 0.01% Triton X-100), 8 mM DTT, and 1.12 U/µL RNase inhibitor. The DNA nanocircle template (NC-45, 80 nM) was added next, followed by reaction initiation with 500 µM NTP solution (each NTP = 500 µM). Aliquots of 2 µL were taken every 30 min for 2 hr starting at t = 0 min and quenched in 75 µL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). Amount of RNA synthesized was measured by the addition of 75 µL of 1:200 diluted RiboGreen in TE buffer, followed by fluorescence measurement using a Synergy H1 Hybrid plate reader with excitation and emission wavelengths of 480 nm and 520 nm, respectively. Each of the compounds was tested in duplicate (n = 2) at eight different concentrations.

An RNA standard curve was made using varying concentrations of tRNA, and the curve was used to convert the fluorescence values to the concentration of RNA at each time point. Data were analyzed and IC<sub>50</sub> values were determined as described in Chapter II, normalizing rate of RNA synthesis to percent activity and fitting data by nonlinear regression.

### **Plasmid-Based Time Course Transcription Assay with Heparin**

Time course transcription assays were prepared as described in Chapter III with the following modifications. The 2x *E. coli* reaction buffer and 2x MTB reaction buffer solutions did not contain NTPs and contained either pMGA12, pMGA4-Ec, or pMGA4-Mtb. The 2x reaction buffer (72  $\mu$ L) was incubated with 75  $\mu$ L of 2x RNAP solution for 20 min. Before initiation with 3  $\mu$ L of 25 mM NTPs, 1.5  $\mu$ L of 5 mg/mL heparin or 1.5  $\mu$ L of H<sub>2</sub>O was added to reactions in triplicate. Aliquots of 10  $\mu$ L were removed from the reaction mixture every 30 min. Before taking the 30-min time point, 1.3  $\mu$ L of 5 mg/mL Heparin was added to a set of reactions (in triplicate) that did not contain heparin already. Fluorescence was measured as described in Chapter III. Heparin was added to the reactions to trap RNAP in an inactive state as depicted in Figure V-3.

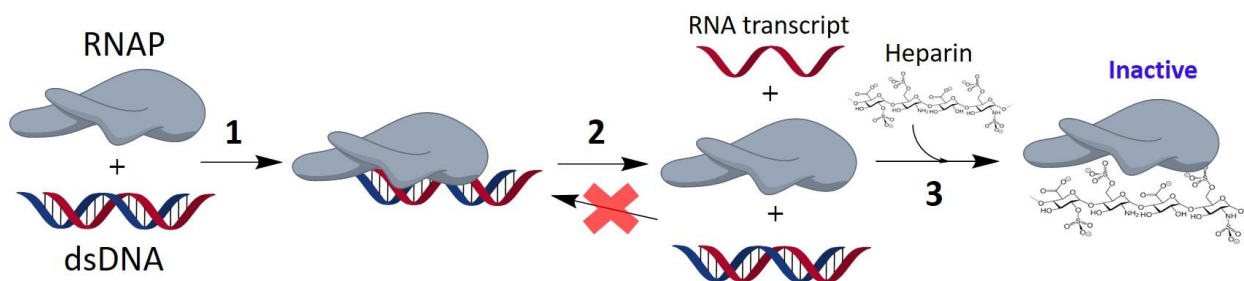


Figure V-3. Schematic of heparin trapping RNAP in an inactive state. 1) Binding of double-stranded DNA (dsDNA) template by RNA polymerase (RNAP) holoenzyme. 2) Synthesis of RNA transcript by RNAP. 3) Binding of Heparin to RNAP that is not bound to dsDNA template. Heparin-bound RNAP is unable to bind dsDNA template and synthesize RNA.

### Plasmid-Based Transcription Assay

Plasmid-based transcription assays were conducted as described in Chapter IV using pMGA4-Ec or pMGA4-Mtb (unless otherwise specified) for *E. coli* and MTB RNAPs, respectively. Compounds were tested in duplicate against WT RNAPs at concentrations of 0.78 – 400 nM with an RNAP concentration of 5 nM ( $\sigma^{70}$ /SigA = 15 nM for *E. coli*/MTB, respectively). Compounds were tested in duplicate against Rif<sup>R</sup> RNAPs (D516V and S531L – *E. coli* numbering) at concentrations of 0.24 – 1000  $\mu$ M (0.24 – 2000  $\mu$ M for rifampin). Rifampin was tested in

triplicate. Data were analyzed and IC<sub>50</sub> values were determined as described in Chapter II, normalizing fluorescence values to percent activity and fitting data by nonlinear regression.

*E. coli* and MTB RNAPs were tested with and without 200 nM rifampin to evaluate endogenous *E. coli* WT RNAP contamination from the preparation (Appendix V-1).

## Results

### Effect of Rifalazil and bxRifs 2b – 2d on *E. coli* RNAPs

The IC<sub>50</sub> values were determined for rifalazil and bxRifs 2b – 2d against *E. coli* WT, D516V, H526Y, and S531L in rolling circle transcription (RiboGreen detection) using core RNAP (Table V-2). The IC<sub>50</sub> values for bxRifs against *E. coli* H526Y RNAP were the highest, with all the values >1000 μM except for rifalazil. The bxRifs 2b – 2d all had IC<sub>50</sub> values < 100 μM against *E. coli* D516V and S531L RNAP.

Table V-2. In vitro IC<sub>50</sub> values (μM) for rifampin, rifalazil, and bxRifs b-d in rolling circle transcription assay

RNAP	Rifampin	Rifalazil	bxRIF2b	bxRIF2c	bxRIF2d
<i>E. coli</i> WT	0.012 <sup>b</sup>	<0.01	0.057	0.024	0.021
<i>E. coli</i> D516V	233 <sup>b</sup>	204	35	2	41
<i>E. coli</i> H526Y	1130 <sup>b</sup>	299	1148	≥ 2000 <sup>a</sup>	1581
<i>E. coli</i> S531L	171 <sup>b</sup>	63	64	92	13
Data for MTB RNAPs					
MTB WT	0.017 <sup>b</sup>	<0.01 <sup>c</sup>	<0.01 <sup>c</sup>	<0.01 <sup>c</sup>	<0.01 <sup>c</sup>
MTB D435V	465 <sup>b</sup>	541 <sup>c</sup>	20 <sup>c</sup>	9 <sup>c</sup>	13 <sup>c</sup>
MTB H445Y	610 <sup>b</sup>	172 <sup>c</sup>	171 <sup>c</sup>	437 <sup>c</sup>	574 <sup>c</sup>
MTB S450L	109 <sup>b</sup>	117 <sup>c</sup>	16 <sup>c</sup>	18 <sup>c</sup>	122 <sup>c</sup>

<sup>a</sup>50% inhibition of transcription was observed at 2000 μM inhibitor

<sup>b</sup>From Gill, et al. (2011) *Tuberculosis* 91, 361-369.

<sup>c</sup>From Gill, et al. (2012) *Journal of Medicinal Chemistry* 55(8), 3814-3826.

X-ray crystal structures were solved for *E. coli* WT RNAP in complex with either rifampin, bxRif2b, or bxRif2c (Appendix V-2). The structure of *E. coli* RNAP with bxRif shows interaction of bxRif C3'-tail with  $\beta$  fork loop 2 and the  $\sigma$  finger of  $\sigma^{70}$  (Figure V-4).

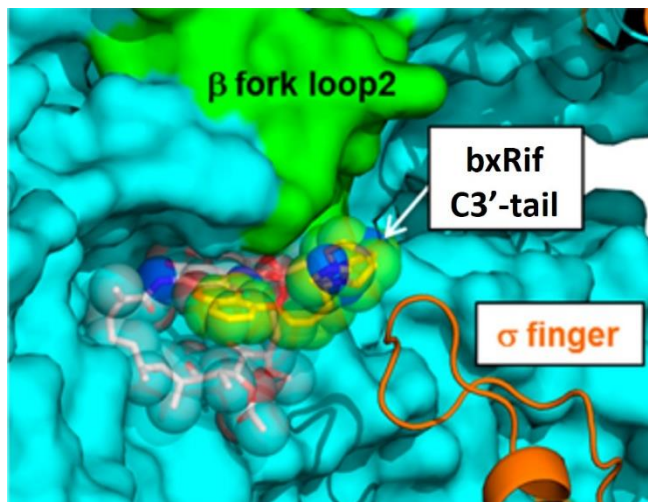


Figure V-4. X-ray crystal structure of bxRif2c bound to *E. coli* RNA polymerase. The structure indicates that the C3'-tail of bxRif2c interacts with both the RNA polymerase and the  $\sigma$  factor.

#### Detection of RNAP Inhibition by Rifampin in Plasmid-Based Transcription Assay

Concentration-response studies were done for rifampin against *E. coli* and MTB RNAPs (WT and Rif<sup>R</sup>) using the plasmid-based transcription assay with pMGA12 or pMGA4-Ec and pMGA4-Mtb, respectively. We were unable to determine any IC<sub>50</sub> values using pMGA12 for *E. coli* Rif<sup>R</sup> RNAPs (D516V, H526Y, and S531L); however, IC<sub>50</sub> values could be determined using the plasmid template pMGA4-Ec (Figure V-5, left). An IC<sub>50</sub> value could be determined for MTB Rif<sup>R</sup> RNAPs using either pMGA12 or pMGA4-Ec, and the concentration-response curves using the two plasmids agreed with one another very well (Figure V-5, right). IC<sub>50</sub> values for rifampin were determined against *E. coli* and MTB RNAPs (WT and Rif<sup>R</sup>, Table V-3) using plasmids pMGA4-Ec and pMGA4-Mtb, respectively. The *E. coli* RNAP IC<sub>50</sub> values correlate well with the MTB RNAPs, with *E. coli* H526Y and MTB H445Y RNAPs being the most resistant to rifampin.



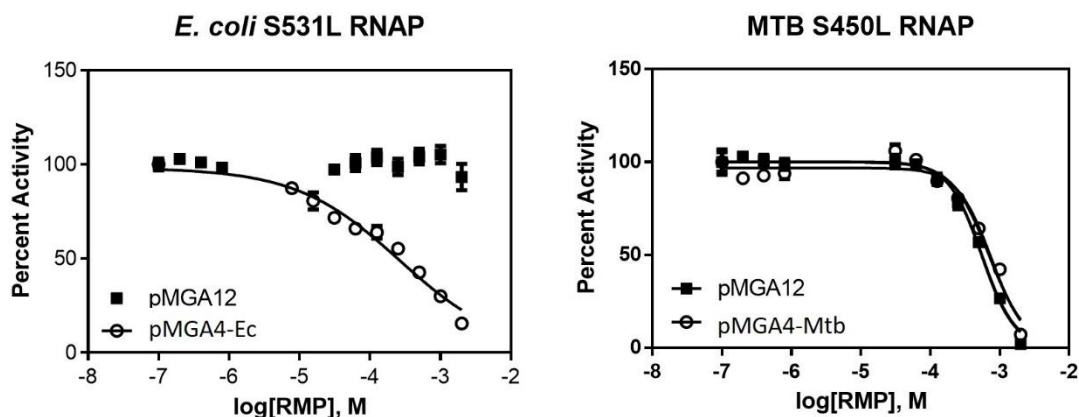


Figure V-5. Concentration-response curves of rifampin against *E. coli* and MTB Rif<sup>R</sup> RNAP using the plasmid-based transcription assay. Plasmid DNA templates used were either pMGA12, pMGA4-Ec, or pMGA4-Mtb with concentration of RNAP and  $\sigma^{70}$ /SigA of 40 and 120 nM, respectively.

Table V-3. RNAP IC<sub>50</sub> values ( $\mu$ M) for rifampin in plasmid-based transcription assay

<i>E. coli</i> RNAPs	WT	D516V	H526Y	S531L
Rifampin	< 0.005	400 ( $\pm$ 120) <sup>a</sup>	$\geq$ 2000 <sup>b</sup>	260 ( $\pm$ 30) <sup>a</sup>
MTB RNAPs	WT	D435V	H445Y	S450L
Rifampin	< 0.005	880 ( $\pm$ 170) <sup>a</sup>	NDI <sup>c</sup>	790 ( $\pm$ 250) <sup>a</sup>

<sup>a</sup>The errors are standard deviation. All experiments were conducted in triplicate.

<sup>b</sup>~50% inhibition at 2 mM, but an IC<sub>50</sub> was not able to be confidently determined.

<sup>c</sup>NDI = no detectable inhibition at 2 mM

### Effect of Heparin on Plasmid-Based Transcription

The activity of *E. coli* and MTB WT RNAPs were monitored in the presence and absence of the negatively charged heparin since it effectively binds and traps RNAP in an inactive state (Figure V-3).<sup>121</sup> The RNAP activity when using the plasmid templates pMGA12 or pMGA4-Ec was monitored over the course of 2 to 3 hours, adding heparin at t = 0 min or t = 30 min (Figure V-6). When using the plasmid template pMGA12, which does not contain any terminator sequences, there is a minimal effect of heparin on the fluorescence detection of transcription for the first 90 min of the reaction (Figure V-6, top left). This indicates that there are very few termination

events where RNAP dissociates from the plasmid DNA template, which would allow heparin to bind DNA and halt transcription. On the other hand, when using pMGA4-Ec, transcription ceases essentially immediately after addition of heparin (Figure V-6, top right).

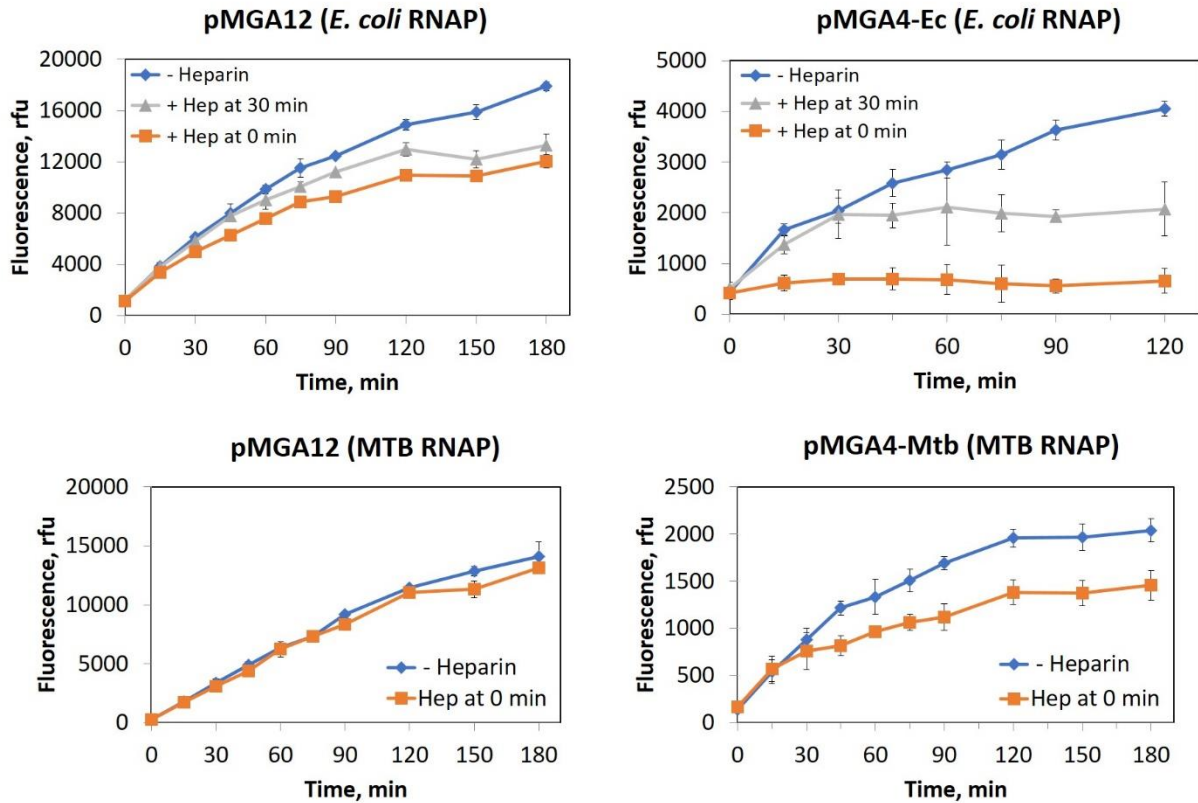


Figure V-6. Time course plasmid-based transcription assay of *E. coli* and MTB WT RNAP with and without heparin. Transcription assays with 40 nM RNAP and 120 nM  $\sigma^{70}$  or SigA. Hep = Heparin, rfu = relative fluorescence units.

A similar observation was seen for MTB RNAP. Heparin has minimal effect on MTB WT RNAP transcription from pMGA12 over the course of the assay (3 hrs). However, the inhibition of transcription by heparin is delayed when MTB RNAP is transcribing from pMGA4-Mtb compared to *E. coli* RNAP transcribing from pMGA4-Ec. It takes approximately 30 min before there is a noticeable reduction in transcription by heparin. This indicates that termination is not as efficient for MTB RNAP with this promoter, or MTB RNAP is much slower at transcribing from the template – as indicated with the low fluorescence values.

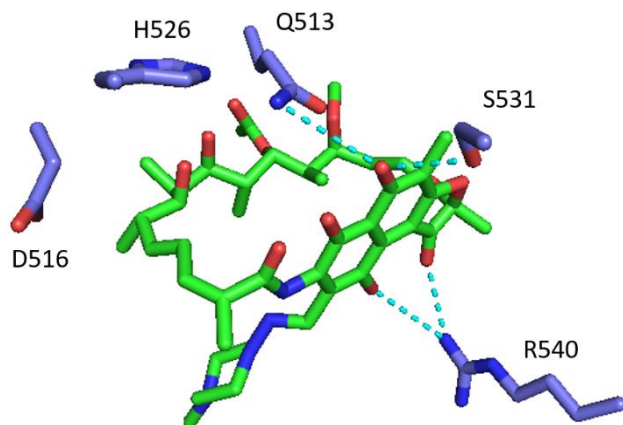


Figure V-7. Interactions of rifampin with residues of *E. coli* RNAP (PDB: 5UAC). Residues in the three most clinically relevant mutations as well as Q513 and R540 are shown.

### Effect of Rif<sup>R</sup> Mutations on *E. coli* RNAPs

In the WT RNAP•RMP complex, the S531 residue forms a hydrogen bond with a phenol of the rifampin naphthalene ring (Figure V-7), and the top of the naphthalene ring is covered by  $\beta$  subunit loops, including fork loop 2 (residues 534-541, Figure V-8, A – left). In the crystal structure of S531L RNAP-rifampin complex; however, a major difference is observed around the  $\beta$  subunit fork loop 2 of which the electron density is weak and scattered compared to its counterpart in the WT RNAP•RMP complex (Figure V-A2). This results in about half of the rifampin naphthalene ring being exposed to solvent (Figure V-8, A – right), reducing approximately 40% of the contact area between RNAP and rifampin (386 Å<sup>2</sup> and 244 Å<sup>2</sup> in WT and S531L, respectively). This dramatically reduces the van der Waals interactions, resulting in a large loss of binding free energy and is consistent with drastic increases in IC<sub>50</sub> of rifampin for the Ser → Leu mutants compared to WT RNAP (Table V-3). Fork loop 2 is ordered in the S531L RNAP without rifampin bound (Figure V-A2), suggesting that the steric clash between rifampin and the Leu side chain causes the observed disorder.

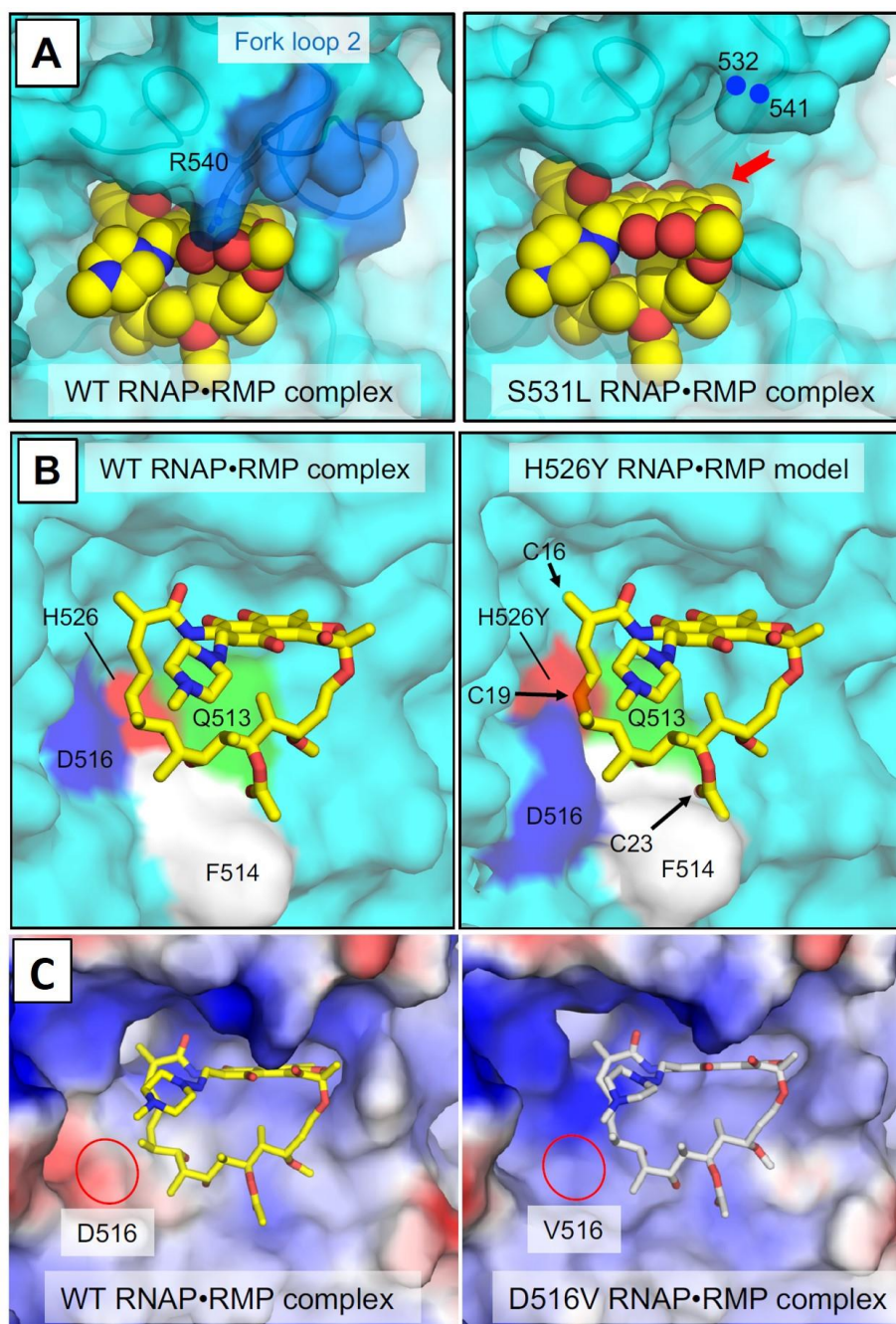


Figure V-8. Structural bases of Rif<sup>R</sup> by clinically relevant mutations S531L, H526Y, and D516V. Top (A): the rifamycin binding pocket of the WT (left) and S531L (right) RNAP complexed with rifampin, shown with the  $\beta$  subunit (cyan, transparent molecular surfaces plus cartoon models) and rifampin (sphere models). Middle (B): WT RNAP•RMP (left) and H526Y RNAP with rifampin overlay (right, based on WT RNAP•RMP complex). Arrows indicate locations of steric clash between ansa bridge of rifampin and H526Y RNAP. Bottom (C): Electrostatic surfaces of rifamycin binding pocket of WT (left) and D516V (right) RNAP complexed with rifampin (stick models). RNAP surfaces are colored with positive (blue), negative (red) and neutral (white) electrostatic potentials. Positions of the D516 and V516 residues are indicated by red circles.

In the WT RNAP•RMP complex, the H526 residue is involved in forming the back wall of the rifamycin binding pocket (Figure V-8, B – left) and may form a hydrogen bond with oxygen atoms of the ansa bridge and/or the Q513 side chain (Figure V-7). The crystal structure of H526Y RNAP was determined at 3.6 Å resolution.<sup>97</sup> A crystal structure with rifampin bound to H526Y RNAP was unable to be determined, presumably due to its poor affinity ( $IC_{50} \geq 2000 \mu M$ ). Modeling rifampin from the WT RNAP•RMP complex into the rifamycin binding pocket of H526Y RNAP structure shows that the rearranged rifamycin binding pocket would sterically clash with the ansa bridge of rifampin (Figure V-8, B – right). This steric clash is consistent with the high level of resistance conferred by the H526Y RNAP in the  $IC_{50}$  values (Table V-3).

In the WT RNAP•RMP complex, the D516 residue is involved in forming part of the sidewall of the rifamycin binding pocket. The D516 residue also contribute to rifampin binding by van der Waals interactions with the ansa bridge of rifampin and/or charge neutralization of basic residues in the rifamycin binding site to facilitate binding of the relatively apolar rifampin. There are no major structural changes in the D516V RNAP mutant, with the RNAP•RMP binding complex similar to WT RNAP (Figure V-8, C and Figure V-A2). The main difference between the two structures is electrostatic distribution of the rifamycin binding pocket, which becomes more basic due to the mutation of Asp→Val – making it less favorable for rifampin binding.

#### **Activity of Novel bxRifs on *E. coli* and MTB RNAPs**

Concentration-response studies were executed using the plasmid-based transcription assay with pMGA4-Ec or pMGA4-Mtb for *E. coli* and MTB RNAPs, respectively. A series of novel bxRifs and RifB analogs were tested to determine  $IC_{50}$  values against *E. coli* and MTB RNAPs (WT, D516V, and S531L RNAPs – *E. coli* numbering). The compounds were clustered into three

different groups: Series A – compounds with an n-butyl linker at position R<sub>2</sub>, Series B – compounds with a benzyl linker at position R<sub>2</sub>, and Series C – compounds with a methyl or formyl group at either position R<sub>1</sub> or R<sub>2</sub>.

Table V-A2 contains IC<sub>50</sub> values for a series of RifB analogs against *E. coli* and MTB RNAPs (WT and Rif<sup>R</sup>). Tables V-A3 thru V-A5 contain IC<sub>50</sub> values for novel bxRifs against *E. coli* and MTB RNAPs. The MICs against replicating (MABA) and non-replicating (LORA) MTB as well as fold activation of hPXR are listed in Tables V-A2 thru V-A5. Rifampin is included at the end of each table for reference. All the compounds tested had similar potency against *E. coli* and MTB WT RNAPs, ranging from < 5 nM to 20 nM. Most of the compounds were more potent against the Asp → Val mutant RNAP than the Ser → Leu mutant RNAP. This was true for both *E. coli* RNAPs and MTB RNAPs. Most of the compounds were more potent against *E. coli* than MTB Rif<sup>R</sup> RNAPs. This is not surprising since the compounds were designed based on X-ray crystal structures of *E. coli* RNAP with bxRifs.

When comparing RifB analogs and bxRif compounds with similar R groups, the RifB analogs were less potent (257824 vs 262463). The IC<sub>50</sub> values for bxRifs against *E. coli* Rif<sup>R</sup> RNAPs ranged from 0.58 μM to > 1000 μM, while for MTB Rif<sup>R</sup> RNAPs they ranged from 7.2 μM to > 1000 μM. The following trends were observed for the bxRif analogs against Rif<sup>R</sup> RNAPs and are outlined in Figure V-9.

- 1) bxRifs with less flexibility (aromatic ring) in linker at position R<sub>2</sub> are more potent
  - 258765 vs 258922 and 262821, 257967 vs 262721, 257983 vs 262806
- 2) bxRifs with larger group (isobutyl vs methyl) on piperazine ring at R<sub>2</sub> are more potent
  - 262803 vs 262805

3) Larger groups on piperazine ring at R<sub>1</sub> are tolerated and potentially more potent against MTB RNAPs

- 262823 vs 262824

4) Carbonyl on piperazine linker at position R<sub>2</sub> has minimal effect on potency

- 257984 vs 257983, 257966 vs 257967

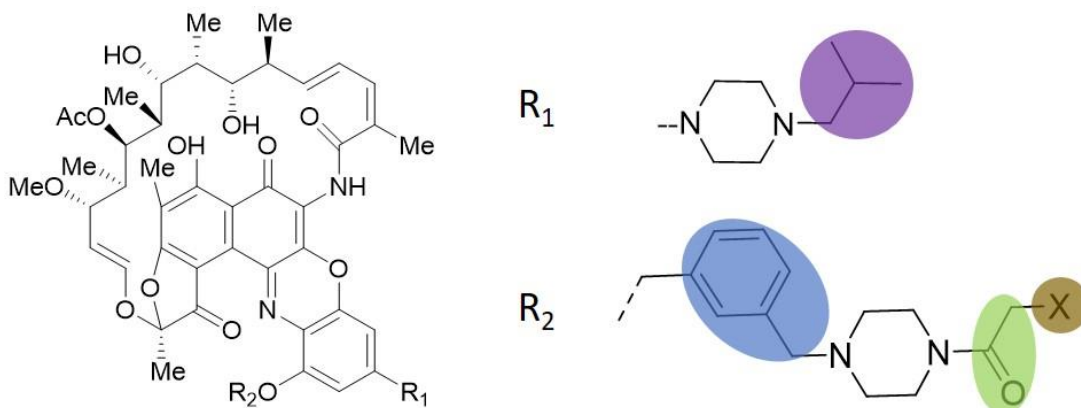


Figure V-9. Overview of bxRif structure-activity relationship studies against Rif<sup>R</sup> RNAP. Purple: Larger substitutions on piperazine are tolerated. Blue: Less flexibility in linker is better for activity. Green: Substitution (carbonyl) on piperazine linker is tolerated. Brown: Various substitutions are tolerated.

The novel bxRifs were tested for cellular activity against MTB in culture, and for activation of hPXR. Rifampin has nanomolar cellular potency against MTB (30 and 330 nM – MABA and LORA, respectively). Rifampin is a strong activator of hPXR, exhibiting a 20 – 30-fold activation of hPXR at concentrations of 20 μM or higher in a eukaryotic cell-based hPXR activation assay. However, several bxRifs disrupt the eukaryotic cell growth at concentrations that high, which affects the reliability of the fold activation in the hPXR assay. To reliably compare hPXR activation, the fold activation at 6.25 μM for each compound is listed in Tables V-A2 thru V-A5, where cell viability is > 90% for all compounds tested (determined via MTT assay). Several compounds were identified that had minimal (less than 2) fold-activation at 6.25 μM and were more potent against MTB in culture than rifampin (262721, 262725, 262823, 262825). It was

observed that compounds with a larger substituent at R<sub>1</sub> (e.g. piperazine) were less likely to activate hPXR. Of the compounds with fold-activation less than 2, all but 262825 have MIC values < 16 nM. It is therefore of interest to test these compounds at lower concentrations to more accurately determine the MICs.

## Discussion

*E. coli* WT and Rif<sup>R</sup> RNAPs have been evaluated in vitro and by X-ray crystallography to evaluate their effect on rifamycin binding and resistance. Studies of bxRifs 2b – 2d inhibitory activity on *E. coli* RNAPs agreed with data previously obtained against MTB, indicating that *E. coli* RNAP is an adequate model for studying the interactions of bxRifs with RNAP and correlating them to MTB RNAP. The crystal structures of *E. coli* WT RNAP with bxRifs confirm the suspected interactions and spatial orientation of the designed bxRifs in the rifamycin binding pocket (Figure V-4) that were designed based on models from *T. thermophilus* RNAP•RMP structure.<sup>57</sup> This confirmation provided additional merit for the development of novel bxRifs that interact with the sigma factor to increase binding affinity.

Previous in vitro evaluations of bxRifs utilized a rolling circle transcription assay that uses a small (45-nt to 90-nt), single-stranded DNA nanocircle.<sup>57</sup> Since this is not the best representation of cellular transcription, we utilized the plasmid-based transcription assay with malachite green detection. However, *E. coli* Rif<sup>R</sup> RNAP in the plasmid-based transcription assay using pMGA12 was unaffected by increasing concentrations of rifampin (Figure V-5). A plasmid containing three synthetic terminator repeats (*synB*)<sup>122</sup> with four repeats of the malachite green aptamer was used (pMGA4-Ec or pMGA4-Mtb) to detect RNAP activity and inhibition by rifamycins. Transcription from this plasmid was susceptible to inhibition by rifampin (Figure V-



5). Since rifamycins inhibit RNA transcription initiation (and not elongation) we postulated that due to the lack of terminator sequences on pMGA12, it is possible for the RNAP to continuously transcribe RNA in the elongation phase without dissociating from the DNA during the 1.5-hr time course of the reaction. This would result in rifampin being unable to inhibit RNAP while RNAP is in the elongation phase.

To test this hypothesis, a time course experiment in the presence and absence of heparin (binds and inactivates RNAP not bound to DNA)<sup>121</sup> was done to compare the rates of RNAP inactivation by heparin for each plasmid template (Figure V-6). It was determined that heparin immediately binds and inactivates *E. coli* RNAP when transcribing from pMGA4-Ec, consistent with frequent transcription termination. However, heparin has minimal effect on RNAP (*E. coli* and MTB) transcription from pMGA12. Even though MTB Rif<sup>R</sup> RNAP transcription from pMGA12 was inhibited by rifampin (Figure V-5, right), the heparin experiment indicated that minimal transcription termination is occurring from pMGA12. We believe that there is no observable inhibition of *E. coli* Rif<sup>R</sup> RNAP transcription by rifampin (when using pMGA12) because *E. coli* RNAP has a fast transcription rate, and the processivity of the enzyme overshadows the inhibition that could be seen if only a single set of repeats were produced. Conversely, inhibition of MTB Rif<sup>R</sup> RNAP transcription from pMGA12 is observable due to the slower rate of transcription, preventing the enzyme from transcribing numerous repeats of plasmid DNA within the time course of the reaction.

With the plasmid-based transcription assay adapted for evaluation of rifamycin inhibition, the effect of *E. coli* Rif<sup>R</sup> RNAPs on rifampin was evaluated by determination of IC<sub>50</sub> values for rifampin and crystal structures of *E. coli* Rif<sup>R</sup> RNAPs with and without rifampin. As previously

reported using rolling circle transcription, H526Y RNAP was the most resistant to rifampin.<sup>56</sup> The contamination from endogenous *E. coli* WT RNAP was also evaluated and determined for *E. coli* and MTB Rif<sup>R</sup> RNAPs (Appendix V-1). The MTB RNAPs were determined to not be contaminated with endogenous *E. coli* WT RNAP. However, *E. coli* Rif<sup>R</sup> RNAPs were ~25% contaminated with endogenous *E. coli* WT RNAP (Figure V-A1). This contamination likely resulted from the Histidine tag being coded for on the *rpoC* gene instead of *rpoB*, where the resistant mutations are, allowing for formation and purification of enzyme with endogenous WT *rpoB* along with Rif<sup>R</sup> RNAP.

Previous studies have used WT RNAP•RMP complex structures (*T. thermophilus*, *T. aquaticus*, more recently *E. coli*) to model the RNAP structures with Rif<sup>R</sup> mutations using molecular dynamics simulations or simple amino acid substitutions to postulate the molecular mechanisms of resistance.<sup>62,117,123,124</sup> The predictions made by Campbell *et al.* that the D516V mutation confers resistance (based on *T. aquaticus* RNAP•RMP structure) via the loss of negative charge are consistent with the *E. coli* D516V RNAP structures. However, previous work on Rif<sup>R</sup> resistance mechanisms did not identify the resistance mechanisms of the S531L and H526Y RNAP supported by the *E. coli* Rif<sup>R</sup> RNAP crystal structures. Based on the crystal structures of S531L RNAP with and without rifampin (Figure V-8, A), resistance occurs not only from elimination of a hydrogen bonding interaction with the naphthalene ring, but also a decrease in van der Waals interactions between the  $\beta$  subunit and the naphthalene ring caused by disordering of fork loop 2. The disordering of fork loop 2 of *E. coli* S531L RNAP was only observed with rifampin bound (Figure V-8, A and Figure V-A2), supporting the steric clash between rifampin and the Leu side chain since structures of *E. coli* S531L RNAP without rifampin did not exhibit any gross structural changes in the rifamycin binding pocket. In contrast, *E. coli* H526Y RNAP had significant changes

in the binding pocket, sterically blocking rifampin binding to RNAP (Figure V-8, B). The H526Y RNAP mutation reshapes the binding site, resulting in insensitivity to rifampin at concentrations up to 2000  $\mu$ M (Table V-3). Based on the structures of the Rif<sup>R</sup> RNAPs, it is not surprising that H526Y RNAP is highly resistant to rifamycins.

Novel bxRifs and RifB analogs were evaluated for activity against WT and Rif<sup>R</sup> RNAPs (*E. coli* and MTB). Overall, the bxRif analogs are more potent than the RifB analogs tested. This is consistent with the higher potency of rifalazil, the bxRif these compounds are based on. However, RifB analogs are still of interest because they are much more synthetically accessible than bxRifs. Most of the compounds were more potent against the Asp  $\rightarrow$  Val mutant RNAP than the Ser  $\rightarrow$  Leu mutant RNAP (for both *E. coli* and MTB). Evaluation of the bxRifs' activities shows that compounds with an aromatic linker at position R<sub>2</sub> are more potent. Decreasing the flexibility of inhibitors can confine them into an active orientation, so it is not surprising that the aromatic linker increases potency relative to a saturated hydrocarbon linker. Modifications to the linker between the piperazine and imidazoline (Series A, Table V-A3) are tolerated. Additional modifications could potentially be made on this linker to decrease flexibility. Most of the novel Rif analogs are more active against *E. coli* RNAPs than MTB RNAPs. This is not surprising since the design of the compounds was based on crystal structures of the *E. coli* WT RNAP. The crystal structure of *M. tuberculosis* RNAP has recently been solved and published.<sup>125</sup> Further design and structural work of bxRifs in complex with bacterial RNAP should be done using the *M. tuberculosis* RNAP to properly evaluate the potential of novel bxRifs as an anti-TB drug.

Four of the novel bxRifs evaluated activated hPXR less than 2-fold (262721, 262725, 262823, 262825) and were more potent than rifampin against MTB in culture. This is very

promising for identification of a rifamycin that does not have the same issues of drug-drug interactions via induction of CYP450s. Additionally, if a more potent rifamycin with good PK/PD properties is identified, this could potentially allow for a shorter treatment duration. It has been shown that higher doses of rifamycins could potentially lead to a shorter duration of therapy; however, there have been reports of tolerability issues for higher doses.<sup>88,126</sup> Therefore, it is reasonable to deduct that a more potent rifamycin analog with similar or better tolerability in patients could reduce the treatment duration.

### Conclusions

X-ray crystal structures of *E. coli* WT RNAP with novel bxRifs and *E. coli* Rif<sup>R</sup> RNAPs with and without rifampin have been determined. These structures show that the designed bxRifs 2b – 2d interact with the sigma factor when bound to *E. coli* RNAP. The *E. coli* Rif<sup>R</sup> RNAP structures indicate structural explanations for the three most clinically relevant mutations conferring rifamycin resistance as follows. In addition to the expected loss of a hydrogen bonding interaction with the naphthalene ring of rifampin, the S531L RNAP mutant results in disordering of  $\beta$  fork loop 2 (when rifampin is bound) causing a loss of van der Waals interactions between fork loop 2 and rifampin. The H526Y RNAP mutant results in rearranging of the rifamycin binding pocket, and overlay of H526Y RNAP and rifampin bound WT RNAP indicates that H526Y RNAP would sterically clash with the ansa bridge of rifampin. This clash prevents rifampin binding and is likely the reason the H526Y RNAP confers such a large degree of resistance on rifampin. As previously predicted based on *T. thermophilus* structures and modelling, the D516V RNAP mutant appears to confer resistance based on the change of electrostatics in the binding pocket resulting from the loss of the negatively charged Asp.

A series of novel bxRifs have been evaluated for activity against *E. coli* and MTB RNAPs (WT and Rif<sup>R</sup>), against *M. tuberculosis* in culture, and for activation of hPXR (leads to CYP450 induction). We have shown that further improvements can be made for activity against RNAPs in vitro and *M. tuberculosis* in culture. However, additional improvements are still necessary against MTB Rif<sup>R</sup> RNAPs to have clinical significance against drug-resistant TB. Further evaluation in PK/PD studies is necessary for the bxRifs that have improved activity against *M. tuberculosis* and do not activate hPXR.

## Notes to Chapter V

### Acknowledgements

I would like to thank Drs. Irosha Nawarathne, Anjanette Turbiak, and Hao Xu for preparation of rifalazil and bxRifs 2b – 2d. I would like to thank Drs. Hollis Showalter and Walajapet Rajeswaran for their synthetic expertise and synthesis of the semi-synthetic benzoxazinorifamycins and rifamycin B analogs. I would like to thank Drs. Scott Franzblau and Baojie Wan (University of Illinois at Chicago) for providing the MIC data against MTB H<sub>37</sub>R<sub>V</sub> strains. I would like to thank Shireen Ashkar for providing the hPXR activation data. I would like to thank Maxwell Stefan for construction of plasmid templates pMGA4-Ec and pMGA4-Mtb and determination of percent active RNAP.

Portions of the work described in this chapter were published in *Journal of Medicinal Chemistry* in 2013 (Volume 56, Issue 11; pages 4758 – 4763) and *Molecular Microbiology* in 2017 (Volume 103, Issue 6, pages 1034 – 1045).

### Abbreviations

bxRif, benzoxazinorifamycin; CYP450, cytochrome P450; DNA, deoxyribonucleic acid; dsDNA, double-stranded DNA; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FPLC, fast protein liquid chromatography; hPXR, human pregnane X receptor; IC<sub>50</sub>, half-maximal inhibitory concentration; LORA, low-oxygen recovery assay; MABA, microplate alamar blue assay; MDR-TB, multidrug-resistant TB; MIC, minimal inhibitory concentration; MTB, *Mycobacterium tuberculosis*; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NTP, nucleoside triphosphate; PK/PD, pharmacokinetic/pharmacodynamic; PMSF, phenylmethylsulfonyl fluoride; RFU, relative fluorescence units; RifB, rifamycin B; Rif<sup>R</sup>, rifampin-resistant; RMP, rifampin; RNA, ribonucleic acid; RNAP, RNA polymerase; TB, tuberculosis; WT, wild-type

## Appendix V-1

### *E. coli* WT R\*NAP Contamination of Rif<sup>R</sup> RNAPs

*E. coli* and MTB Rif<sup>R</sup> RNAPs were tested in the plasmid-based transcription assay with and without 200 nM rifampin to evaluate endogenous *E. coli* WT RNAP contamination from the preparation. MTB Rif<sup>R</sup> RNAPs were not contaminated with endogenous *E. coli* WT RNAP. The total active RNAP was determined using a radiometric single turnover assay (Max Stefan). The specific activity of WT RNAP was determined by dividing the fluorescence reading by the active RNAP concentration. The amount of contaminating endogenous WT RNAP was determined for each Rif<sup>R</sup> mutant, then subtracted from total active RNAP. The percent Rif<sup>R</sup> mutant in each prep determined from the calculations (Table V-A1) are shown in Figure V-A1.

Table V-A1. Determination of endogenous *E. coli* WT RNAP in Rif<sup>R</sup> mutant purifications

<i>E. coli</i> RNAPs	WT	D516V	H526Y	S531L
Active RNAP (%)	100 ± 4	104 ± 3	75 ± 3	64 ± 1
[Total RNAP] <sub>Active</sub> (nM)	20	20.8	15	12.8
Fluorescence, - RMP (RFU)	14823	8460	4215.5	7089
Specific Activity (RFU/nM)	741			
Fluorescence, WT contaminant (RFU)	N/A	3557	2890	1704
[WT RNAP] <sub>contaminant</sub> (nM)	N/A	4.8	3.9	2.3
Percent Rif <sup>R</sup> RNAP	N/A	77%	74%	82%

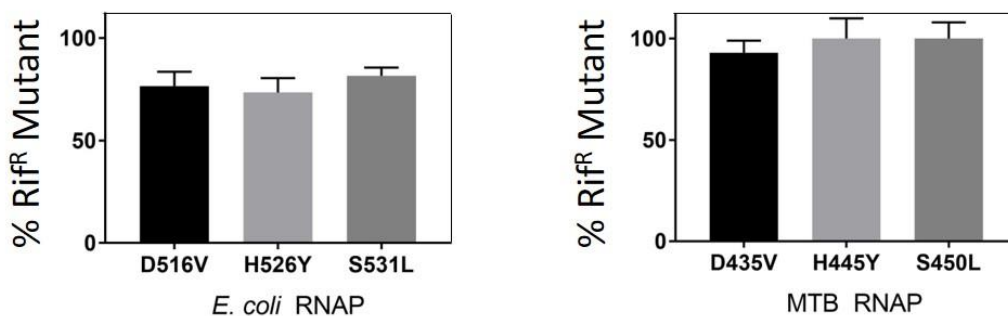


Figure V-A1. Percent Rif<sup>R</sup> mutant in *E. coli* and MTB RNAP Purifications. Left: Percent Rif<sup>R</sup> mutants for *E. coli* RNAPs. Right: Percent Rif<sup>R</sup> mutants for MTB RNAPs.

## Appendix V-2

### X-Ray Crystal Structures of *E. coli* RNAPs

Crystallization studies were performed by Drs. Vadim Molodtsov and Katsu Murakami as described.<sup>97</sup>

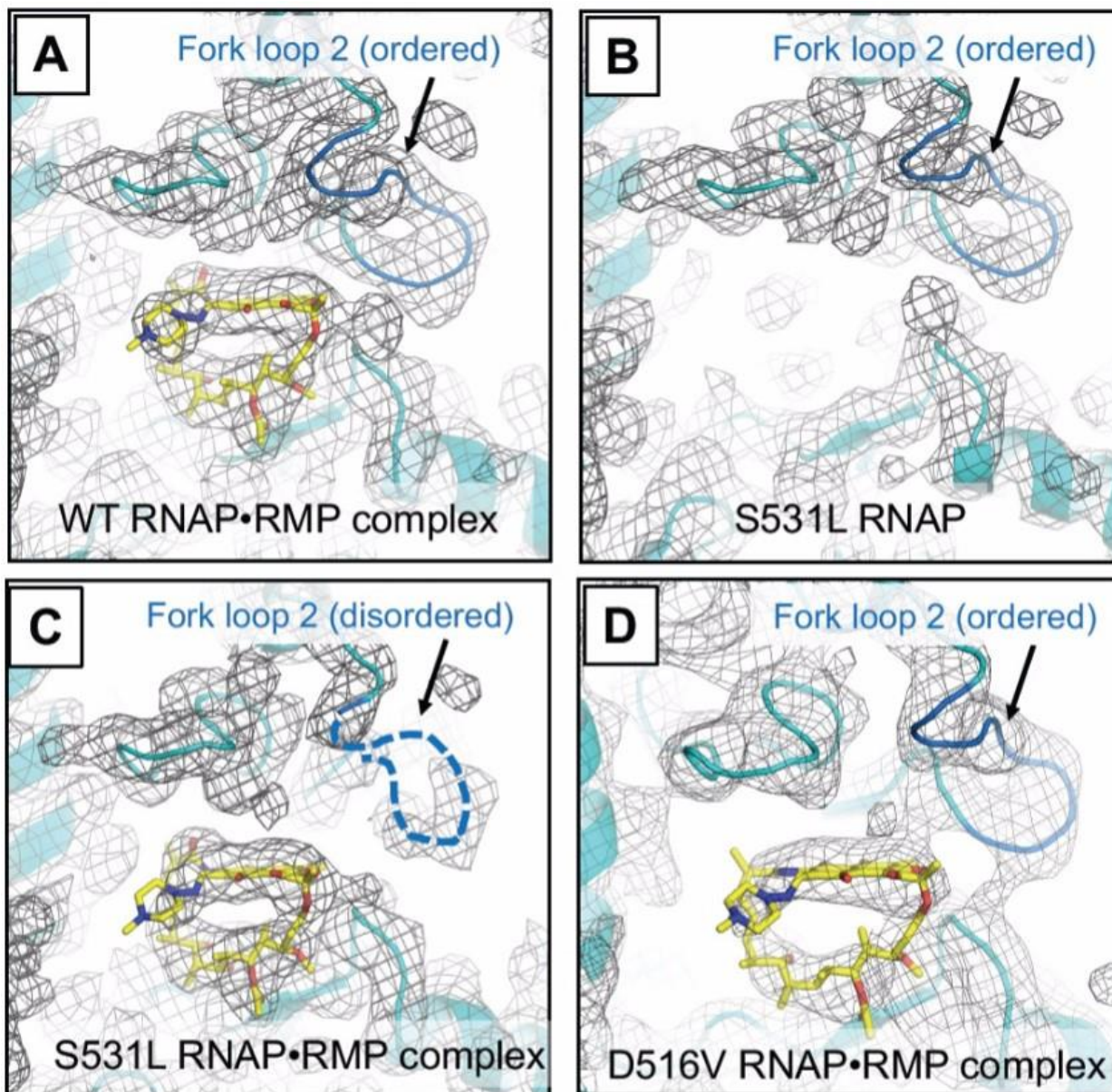


Figure V-A2. Electron density maps of RNAP and RNAP•RMP complexes.  $2F_o - F_c$  electron density map (gray mesh,  $\sigma$  cutoff = 1) of WT RNAP•RMP complex (A), S531L RNAP (B), S531L RNAP•RMP complex (C) and D516V RNAP•RMP complex (D). The subunit and rifampin are shown as cartoon and stick models, respectively. Fork loop 2 of  $\beta$  subunit is colored in blue and labeled. The disordered fork loop 2 in the S531L RNAP•RMP complex is indicated as a dashed line.



## Appendix V-3

### IC<sub>50</sub> Values of bxRifs and RifB Analogs

IC<sub>50</sub> values were determined against *E. coli* and MTB RNAPs using the plasmid-based transcription assay. WT enzymes were tested at 5 nM RNAP, 15 nM  $\sigma^{70}$ /SigA, and 10 nM pMGA4-Ec or pMGA4-Mtb. Rif<sup>R</sup> RNAPs were tested at 40 nM RNAP, 120 nM  $\sigma^{70}$ /SigA, and 80 nM pMGA4-Ec or 40 nM pMGA4-Mtb. Compounds were tested at concentrations of 0.78 – 400 nM for WT RNAPs and 0.24 – 1000  $\mu$ M for Rif<sup>R</sup> RNAPs.

Table V-A2. In vitro activity of RifB analogs against *E. coli* and MTB RNAPs

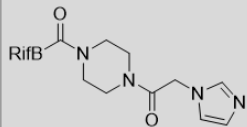
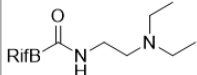
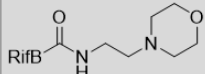
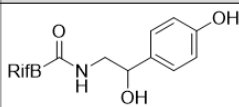
Compound	R	<i>E. coli</i> RNAP IC <sub>50</sub> ( $\mu$ M)			MTB RNAP IC <sub>50</sub> ( $\mu$ M)			MIC <sub>90</sub> (nM)		hPXR, Fold Act. at 6.25 $\mu$ M
		WT	D516V	S531L	WT	D435V	S450L	MABA	LORA	
258450		16 nM	52	320	< 5 nM	380	260	N/A	N/A	5.4
258766		5 nM	150	> 1000	< 5 nM	> 1000	> 1000	N/A	N/A	1.8
262463		8 nM	140	910	5 nM	> 1000	780	460	2900	2.4
262465		5 nM	590	830	10 nM	470	200	720	7100	4.1
Rifampin		< 5 nM	400	260	< 5 nM	880	800	30	330	16

Table V-A3. In vitro activity of benzoxazinorifamycins Series A against *E. coli* and MTB RNAPs

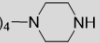
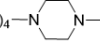
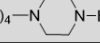
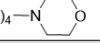
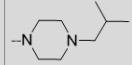
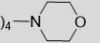
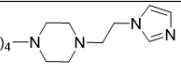
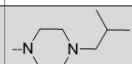
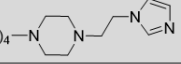
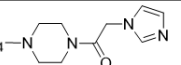
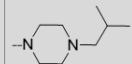
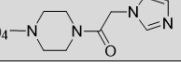
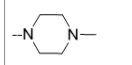
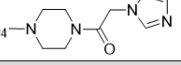
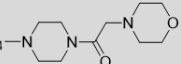
Compound	R <sub>1</sub>	R <sub>2</sub>	<i>E. coli</i> RNAP IC <sub>50</sub> (μM)			MTB RNAP IC <sub>50</sub> (μM)			MIC <sub>90</sub> (nM)		hPXR, Fold Act. at 6.25 μM
			WT	D516V	S531L	WT	D435V	S450L	MABA	LORA	
257822	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	6 nM	29	440	< 5 nM	190	380	N/A	N/A	6.4
257823	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	6 nM	27	790	< 5 nM	400	610	92	190	5.3
257825	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	< 5 nM	22	300	< 5 nM	190	200	24	58	N/A
257824	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	7 nM	65	390	< 5 nM	240	580	21	50	3.3
257965		-(CH <sub>2</sub> ) <sub>4</sub> -N 	6 nM	39	80	< 5 nM	140	160	24	23	3.3
257984	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	7 nM	300	> 1000	15 nM	530	780	50	80	6.1
257966		-(CH <sub>2</sub> ) <sub>4</sub> -N 	6 nM	51	52	7 nM	110	170	47	15	7.6
257983	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	20 nM	260	> 500	10 nM	> 500	> 500	90	190	4.1
257967		-(CH <sub>2</sub> ) <sub>4</sub> -N 	8 nM	59	93	6 nM	160	240	39	21	2.2
258449		-(CH <sub>2</sub> ) <sub>4</sub> -N 	10 nM	12	230	< 5 nM	140	420	82	47	2.7
258765	H	-(CH <sub>2</sub> ) <sub>4</sub> -N 	7 nM	23	> 1000	< 5 nM	240	1000	60	60	5.4
Rifampin			< 5 nM	400	260	< 5 nM	880	800	30	330	16

Table V-A4. In vitro activity of benzoxazinorifamycins Series B against *E. coli* and MTB RNAPs

Compound	R <sub>1</sub>	R <sub>2</sub>	<i>E. coli</i> RNAP IC <sub>50</sub> (μM)			MTB RNAP IC <sub>50</sub> (μM)			MIC <sub>90</sub> (nM)		hPXR, Fold Act. at 6.25 μM
			WT	D516V	S531L	WT	D435V	S450L	MABA	LORA	
258923	H		19 nM	8.8	330	8 nM	61	640	60	120	7.3
262466	H		< 5 nM	0.9	94	< 5 nM	21	180	< 100	< 100	3.3
262464	H		< 5 nM	0.58	28	7 nM	7.2	94	< 100	130	3.1
258922	H		10 nM	39	>500	8 nM	8.8	400	60	90	N/A
262821	H		5 nM	10	92	5 nM	43	150	< 16	30	2.3
262725			16 nM	7.3	32	6 nM	57	110	< 16	< 16	1.5
262721			14 nM	17	66	5 nM	20	150	< 16	< 16	1.1
262806	H		5 nM	24	160	5 nM	53	220	30	60	2.5
262722			20 nM	12	100	< 5 nM	44	74	20	30	1.4
Rifampin			< 5 nM	400	260	< 5 nM	880	800	30	330	16

Table V-A5. In vitro activity of benzoxazinorifamycins Series C against *E. coli* and MTB RNAPs

Compound	R <sub>1</sub>	R <sub>2</sub>	<i>E. coli</i> RNAP IC <sub>50</sub> (μM)			MTB RNAP IC <sub>50</sub> (μM)			MIC <sub>90</sub> (nM)		hPXR, Fold Act. at 6.25 μM
			WT	D516V	S531L	WT	D435V	S450L	MABA	LORA	
262801	CH <sub>3</sub>	H	5 nM	19	46	15 nM	120	310	< 16	< 16	15.1
262805	CH <sub>3</sub>		7 nM	32	350	< 5 nM	14	410	< 16	< 16	4.6
262825			10 nM	30	330	6 nM	260	320	50	120	1.3
262803	CH <sub>3</sub>		6 nM	10	130	< 5 nM	25	200	< 16	< 16	9.7
262822	H	CH <sub>3</sub>	5 nM	38	82	< 5 nM	59	220	< 16	< 16	7.5
262824		CH <sub>3</sub>	6 nM	6.4	46	< 5 nM	35	240	< 16	< 16	5.2
262823		CH <sub>3</sub>	5 nM	12	52	6 nM	24	63	< 16	< 16	2
Rifampin			< 5 nM	400	260	< 5 nM	880	800	30	330	16

## Chapter VI

### Summary, Conclusions, and Future Directions

Tuberculosis (TB) is an infectious disease caused by the pathogen *Mycobacterium tuberculosis*, and is one of the largest global health problems, infecting an estimated 8 – 10 million new people each year.<sup>2</sup> The issue of drug-resistant TB has resulted in a growing global concern due to the prevalence of resistant strains and the limited number of novel anti-TB drugs in development.<sup>127,128</sup> Reinvestigation of bacterial RNA polymerase for drug-resistant TB is of interest since introduction of rifampin in the 1960s led to a reduction in therapy from 18 months to 9 months, with a current treatment regimen of 6 months with rifampin and three other drugs.<sup>43</sup> This dissertation focused on studying novel inhibitors of bacterial RNA polymerase (RNAP) for their potential as inhibitors of *M. tuberculosis* wild-type (WT) and rifampin-resistant (Rif<sup>R</sup>) RNAP.

In chapter II, a plasmid-based transcription assay was developed and optimized using an RNA aptamer that binds malachite green, resulting in an increase in fluorescence of malachite green. This assay utilizes a plasmid template, is cost effective, and is amenable to high-throughput screening. It was shown to be able to detect concentration-dependent inhibition of *E. coli* RNAP by rifampin.

In chapter III, the plasmid-based transcription assay was optimized and implemented for a high-throughput screen against the *E. coli* RNAP analog ( $\beta$ S531L/ $\beta$ 'V408G) of the most clinically relevant *M. tuberculosis* rifampin-resistant (Rif<sup>R</sup>) RNAP. This RNAP contains a mutation that

confers rifamycin resistance ( $\beta$ S531L) as well as a compensatory mutation ( $\beta$ V408G) that is believed to ameliorate a fitness defect.<sup>89</sup> The high-throughput screen resulted in identification of 15 compounds with IC<sub>50</sub> values against *E. coli* or MTB RNAP less than 100  $\mu$ M.

In chapter IV, the top scaffolds identified from the high-throughput screen were evaluated for their activity against bacterial RNAP, *M. tuberculosis* in culture, and their potential for further development against bacterial RNAP. Several scaffolds of interest were evaluated by an initial structure-activity relationship study with commercially available analogs. A cyanopyrimidine scaffold was identified as a specific inhibitor of bacterial RNAP with the potential to be further optimized. It inhibited both WT and Rif<sup>R</sup> RNAP (*E. coli* and MTB), and was evaluated for potential pan-assay interference that could be caused by an aromatic nitrile. The scaffold was found to reversibly inhibit RNAP and does not form an adduct with cysteine in a nitrile reactivity assay, which has previously been used to evaluate the electrophilicity of aromatic nitriles.<sup>100</sup>

In chapter V, we determined structures of *E. coli* WT RNAP with bxRifs that confirmed interactions between bxRifs and the sigma factor, providing further merit to pursue optimization of bxRifs for improved binding and potency. We evaluated the effects of Rif<sup>R</sup> mutations on RNAP as well as their effects on rifampin binding and inhibition of RNAP. Structures of *E. coli* Rif<sup>R</sup> RNAPs with and without rifampin identified resistance mechanisms beyond previous modeling and predictions made from wild-type bacterial RNAPs (*T. thermophilus*, *T. aquaticus*, *E. coli*). In addition, we developed and evaluated novel bxRifs designed based on crystal structures of *E. coli* RNAP in complex with bxRifs. We identified novel bxRifs that are more potent against *M. tuberculosis* in culture than rifampin and show minimal activation of hPXR.

In conclusion, we have identified novel compounds (cyanopyrimidines) that are able to inhibit both WT and Rif<sup>R</sup> RNAP (*E. coli* and MTB). These compounds have modest activity in vitro (low  $\mu\text{M}$  at best) and further structure-activity studies are necessary to identify more potent analogs. Identification of the binding site of this scaffold would allow for structure-based design of cyanopyrimidines with improved activity. X-ray crystal structures of RNAP in complex with an active cyanopyrimidine, followed by mutagenesis studies to confirm binding site would aid in the development of this scaffold as a potent RNAP inhibitor.

We have also developed novel bxRifs with modest improvements in activity against Rif<sup>R</sup> RNAPs. Further, several bxRifs developed are exceedingly potent against *M. tuberculosis* in culture (< 16 nM) and show minimal activation of the human pregnane X receptor, which leads to CYP450 induction and drug-drug interactions in patients. These compounds warrant further work to evaluate their pharmacokinetic and pharmacodynamic properties, followed by studies in mice.

To further aid in the fight against TB and drug-resistance, as well as working towards the eradication of TB, significant progress is still needed. The current 6-month treatment for drug-sensitive TB still presents difficulties in adherence throughout the entire regimen. Research has indicated that higher concentrations of rifampin, if tolerated, could reduce the treatment regimen.<sup>126</sup> If modifications to the current treatment regimen resulted in a shorter duration of treatment, this could help decrease the burden of drug administration and adherence. The long treatment duration results in difficulties not only in adherence, but also administration.

Since 1994, The World Health Organization has recommended a direct observed therapy, short-course (DOTS) for the treatment and eradication of TB; where hospital workers or trained

volunteers directly observe and record the swallowing of pills.<sup>129</sup> This strategy outlined a framework for effective identification and treatment of TB. This framework has been expanded upon to include MDR-TB treatment; however, there are still issues with effective administration and treatment in countries with impoverished health systems.<sup>129</sup> In countries where some of the main tenants of DOTS (government commitment, steady drug supply, standardized recording and reporting) are compromised, it is important to alleviate and simplify these issues. The high cost of second-line TB drugs for MDR-TB and the inability to directly observe and record treatment are both hurdles in remote areas with impoverished health systems. The implementation of a treatment administration strategy that that allows for the successful completion of TB treatment, even in remote areas, will help aid in the fight to eradicate TB.

## References

1. World Health Organization. *Global Health Estimates 2015*. (World Health Organization, 2015).
2. World Health Organization. *Global Tuberculosis Report 2016*. (World Health Organization, 2016).
3. Davies, J. Where have All the Antibiotics Gone? *Can. J. Infect. Dis. Med. Microbiol. J. Can. Mal. Infect. Microbiol. Medicale* **17**, 287–290 (2006).
4. Centers for Disease Control and Prevention, O. of I. D. *Antibiotic resistance threats in the United States*. (Centers for Disease Control and Prevention, Office of Infectious Disease, 2013).
5. Federal Drug Administration. *FDA Approved Drugs, Center for Drug Evaluation and Research*. (Federal Drug Administration, 2017).
6. Iredell, J., Brown, J. & Tagg, K. Antibiotic resistance in Enterobacteriaceae: mechanisms and clinical implications. *BMJ* h6420 (2016). doi:10.1136/bmj.h6420
7. Spellberg, B. *et al.* The Epidemic of Antibiotic-Resistant Infections: A Call to Action for the Medical Community from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **46**, 155–164 (2008).
8. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T Peer-Rev. J. Formul. Manag.* **40**, 277–283 (2015).
9. Coates, A. R., Halls, G. & Hu, Y. Novel classes of antibiotics or more of the same?: New antibiotic classes are urgently needed. *Br. J. Pharmacol.* **163**, 184–194 (2011).
10. Conly, J. & Johnston, B. Where are all the new antibiotics? The new antibiotic paradox. *Can. J. Infect. Dis. Med. Microbiol. J. Can. Mal. Infect. Microbiol. Medicale* **16**, 159–160 (2005).
11. Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13**, 42–51 (2015).
12. Dever, L. A. & Dermody, T. S. Mechanisms of bacterial resistance to antibiotics. *Arch. Intern. Med.* **151**, 886–895 (1991).



13. Morens, D. M., Folkers, G. K. & Fauci, A. S. The challenge of emerging and re-emerging infectious diseases. *Nature* **430**, 242–249 (2004).
14. Tuberculosis: a global emergency. *World Health Forum* **14**, 438 (1993).
15. Barry, C. E. *et al.* The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* (2009). doi:10.1038/nrmicro2236
16. Lin, P. L. & Flynn, J. L. Understanding Latent Tuberculosis: A Moving Target. *J. Immunol.* **185**, 15–22 (2010).
17. Song, S.-E. *et al.* Comparison of the Tuberculin Skin Test and Interferon Gamma Release Assay for the Screening of Tuberculosis in Adolescents in Close Contact with Tuberculosis TB Patients. *PLoS ONE* **9**, e100267 (2014).
18. World Health Organization. *Systematic screening for active tuberculosis: principles and recommendations.* (World Health Organization, 2013).
19. Vynnycky, E. & Fine, P. E. Lifetime risks, incubation period, and serial interval of tuberculosis. *Am. J. Epidemiol.* **152**, 247–263 (2000).
20. Basso, L. A., Zheng, R., Musser, J. M., Jacobs, W. R. & Blanchard, J. S. Mechanisms of isoniazid resistance in Mycobacterium tuberculosis: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. *J. Infect. Dis.* **178**, 769–775 (1998).
21. Sensi, P. History of the development of rifampin. *Rev. Infect. Dis.* **5 Suppl 3**, S402-406 (1983).
22. Takayama, K. & Kilburn, J. O. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrob. Agents Chemother.* **33**, 1493–1499 (1989).
23. Zhang, Y. Mode of action of pyrazinamide: disruption of Mycobacterium tuberculosis membrane transport and energetics by pyrazinoic acid. *J. Antimicrob. Chemother.* **52**, 790–795 (2003).
24. Connolly, L. E., Edelstein, P. H. & Ramakrishnan, L. Why Is Long-Term Therapy Required to Cure Tuberculosis? *PLoS Med.* **4**, e120 (2007).
25. 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. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* **3**, S231-279 (1999).
26. Hooper, D. C. Mode of action of fluoroquinolones. *Drugs* **58 Suppl 2**, 6–10 (1999).

27. Kotra, L. P., Haddad, J. & Mobashery, S. Aminoglycosides: Perspectives on Mechanisms of Action and Resistance and Strategies to Counter Resistance. *Antimicrob. Agents Chemother.* **44**, 3249–3256 (2000).
28. Koul, A. *et al.* Diarylquinolines target subunit c of mycobacterial ATP synthase. *Nat. Chem. Biol.* **3**, 323–324 (2007).
29. Lambert, M. P. & Neuhaus, F. C. Mechanism of D-cycloserine action: alanine racemase from *Escherichia coli* W. *J. Bacteriol.* **110**, 978–987 (1972).
30. Lin, A. H., Murray, R. W., Vidmar, T. J. & Marotti, K. R. The oxazolidinone eperzolid binds to the 50S ribosomal subunit and competes with binding of chloramphenicol and lincomycin. *Antimicrob. Agents Chemother.* **41**, 2127–2131 (1997).
31. Matsumoto, M. *et al.* OPC-67683, a Nitro-Dihydro-Imidazooxazole Derivative with Promising Action against Tuberculosis In Vitro and In Mice. *PLoS Med.* **3**, e466 (2006).
32. Tsukamura, M., Tsukamura, S., Mizuno, S. & Nakano, E. THE MECHANISM OF ACTION OF ETHIONAMIDE. *Am. Rev. Respir. Dis.* **89**, 933–935 (1964).
33. Zheng, J. *et al.* para-Aminosalicylic Acid Is a Prodrug Targeting Dihydrofolate Reductase in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **288**, 23447–23456 (2013).
34. Mahajan, R. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. *Int. J. Appl. Basic Med. Res.* **3**, 1 (2013).
35. Bloemberg, G. V. *et al.* Acquired Resistance to Bedaquiline and Delamanid in Therapy for Tuberculosis. *N. Engl. J. Med.* **373**, 1986–1988 (2015).
36. Hoffmann, H. *et al.* Delamanid and Bedaquiline Resistance in *Mycobacterium tuberculosis* Ancestral Beijing Genotype Causing Extensively Drug-Resistant Tuberculosis in a Tibetan Refugee. *Am. J. Respir. Crit. Care Med.* **193**, 337–340 (2016).
37. Tadolini, M. *et al.* First case of extensively drug-resistant tuberculosis treated with both delamanid and bedaquiline. *Eur. Respir. J.* **48**, 935–938 (2016).
38. Mukhtar, T. A. & Wright, G. D. Streptogramins, oxazolidinones, and other inhibitors of bacterial protein synthesis. *Chem. Rev.* **105**, 529–542 (2005).
39. Zervosen, A., Sauvage, E., Frère, J.-M., Charlier, P. & Luxen, A. Development of new drugs for an old target: the penicillin binding proteins. *Mol. Basel Switz.* **17**, 12478–12505 (2012).
40. Bala, S. *et al.* Reclassification of *Amycolatopsis mediterranei* DSM 46095 as *Amycolatopsis rifamycinica* sp. nov. *Int. J. Syst. Evol. Microbiol.* **54**, 1145–1149 (2004).

41. Füresz, S. & Timbal, M. T. Antibacterial Activity of Rifamycins. *Chemotherapy* **7**, 200–208 (1963).
42. Bergamini, N. & Fowst, G. Rifamycin SV. A review. *Arzneimittelforschung*. **15**, Suppl:951-1002 (1965).
43. Binda, G. *et al.* Rifampicin, a general review. *Arzneimittelforschung*. **21**, 1907–1977 (1971).
44. Portero, J.-L. & Rubio, M. New anti-tuberculosis therapies. *Expert Opin. Ther. Pat.* **17**, 617–637 (2007).
45. Floss, H. G. & Yu, T.-W. Rifamycin-mode of action, resistance, and biosynthesis. *Chem. Rev.* **105**, 621–632 (2005).
46. Aristoff, P. A., Garcia, G. A., Kirchoff, P. D. & Showalter, H. D. Rifamycins--obstacles and opportunities. *Tuberc. Edinb. Scotl.* **90**, 94–118 (2010).
47. Lehmann, J. M. *et al.* The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *J. Clin. Invest.* **102**, 1016–1023 (1998).
48. Saito, H. *et al.* In vitro antimycobacterial activities of newly synthesized benzoxazinorifamycins. *Antimicrob. Agents Chemother.* **35**, 542–547 (1991).
49. Mae, T. *et al.* Effect of a new rifamycin derivative, rifalazil, on liver microsomal enzyme induction in rat and dog. *Xenobiotica Fate Foreign Compd. Biol. Syst.* **28**, 759–766 (1998).
50. Moghazeh, S. L. *et al.* Comparative antimycobacterial activities of rifampin, rifapentine, and KRM-1648 against a collection of rifampin-resistant Mycobacterium tuberculosis isolates with known rpoB mutations. *Antimicrob. Agents Chemother.* **40**, 2655–2657 (1996).
51. Yang, B. *et al.* Relationship between antimycobacterial activities of rifampicin, rifabutin and KRM-1648 and rpoB mutations of Mycobacterium tuberculosis. *J. Antimicrob. Chemother.* **42**, 621–628 (1998).
52. Dietze, R. *et al.* Safety and bactericidal activity of rifalazil in patients with pulmonary tuberculosis. *Antimicrob. Agents Chemother.* **45**, 1972–1976 (2001).
53. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* **79**, 3–29 (1998).
54. Telenti, A. *et al.* Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis. *Lancet Lond. Engl.* **341**, 647–650 (1993).

55. Williams, D. L. *et al.* Contribution of *rpoB* mutations to development of rifamycin cross-resistance in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **42**, 1853–1857 (1998).
56. Gill, S. K. & Garcia, G. A. Rifamycin inhibition of WT and Rif-resistant *Mycobacterium tuberculosis* and *Escherichia coli* RNA polymerases in vitro. *Tuberc. Edinb. Scotl.* **91**, 361–369 (2011).
57. Gill, S. K. *et al.* Structure-based design of novel benzoxazinorifamycins with potent binding affinity to wild-type and rifampin-resistant mutant *Mycobacterium tuberculosis* RNA polymerases. *J. Med. Chem.* **55**, 3814–3826 (2012).
58. *Antibiotics: targets, mechanisms and resistance.* (Wiley-VCH-Verl, 2014).
59. André, E. *et al.* Novel synthetic molecules targeting the bacterial RNA polymerase assembly. *J. Antimicrob. Chemother.* **57**, 245–251 (2006).
60. Artsimovitch, I., Chu, C., Lynch, A. S. & Landick, R. A new class of bacterial RNA polymerase inhibitor affects nucleotide addition. *Science* **302**, 650–654 (2003).
61. Bayro, M. J. *et al.* Structure of Antibacterial Peptide Microcin J25: A 21-Residue Lariat Protoknot. *J. Am. Chem. Soc.* **125**, 12382–12383 (2003).
62. Campbell, E. A. *et al.* Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674–682 (2005).
63. Ciciliato, I. *et al.* Antibiotics GE23077, novel inhibitors of bacterial RNA polymerase. I. Taxonomy, isolation and characterization. *J. Antibiot. (Tokyo)* **57**, 210–217 (2004).
64. Degen, D. *et al.* Transcription inhibition by the depsipeptide antibiotic salinamide A. *eLife* **3**, (2014).
65. Irschik, H., Gerth, K., Höfle, G., Kohl, W. & Reichenbach, H. The myxopyronins, new inhibitors of bacterial RNA synthesis from *Myxococcus fulvus* (Myxobacterales). *J. Antibiot. (Tokyo)* **36**, 1651–1658 (1983).
66. Siddhikol, C., Erbstoesz, J. W. & Weisblum, B. Mode of action of streptolydigin. *J. Bacteriol.* **99**, 151–155 (1969).
67. Sonenshein, A. L. & Alexander, H. B. Initiation of transcription in vitro inhibited by lipiarmycin. *J. Mol. Biol.* **127**, 55–72 (1979).
68. Moore, B. S. *et al.* Salinamides, Antiinflammatory Depsipeptides from a Marine Streptomyces. *J. Org. Chem.* **64**, 1145–1150 (1999).

69. André, E. *et al.* A multiwell assay to isolate compounds inhibiting the assembly of the prokaryotic RNA polymerase. *Assay Drug Dev. Technol.* **2**, 629–635 (2004).
70. Bhat, J. *et al.* High-throughput screening of RNA polymerase inhibitors using a fluorescent UTP analog. *J. Biomol. Screen.* **11**, 968–976 (2006).
71. Babendure, J. R., Adams, S. R. & Tsien, R. Y. Aptamers switch on fluorescence of triphenylmethane dyes. *J. Am. Chem. Soc.* **125**, 14716–14717 (2003).
72. Furukawa, K. *et al.* Fluorescence generation from tandem repeats of a malachite green RNA aptamer using rolling circle transcription. *Bioorg. Med. Chem. Lett.* **18**, 4562–4565 (2008).
73. Daubendiek, S. L. & Kool, E. T. Generation of catalytic RNAs by rolling transcription of synthetic DNA nanocircles. *Nat. Biotechnol.* **15**, 273–277 (1997).
74. Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R. & Darst, S. A. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* **4**, (2015).
75. Ilgu, M. *et al.* Light-up and FRET aptamer reporters; evaluating their applications for imaging transcription in eukaryotic cells. *Methods San Diego Calif* **98**, 26–33 (2016).
76. Chauhan, S. & Tyagi, J. S. Cooperative binding of phosphorylated DevR to upstream sites is necessary and sufficient for activation of the Rv3134c-devRS operon in Mycobacterium tuberculosis: implication in the induction of DevR target genes. *J. Bacteriol.* **190**, 4301–4312 (2008).
77. Wang, W. & Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *BioTechniques* **26**, 680–682 (1999).
78. Brewster, R. C., Jones, D. L. & Phillips, R. Tuning promoter strength through RNA polymerase binding site design in Escherichia coli. *PLoS Comput. Biol.* **8**, e1002811 (2012).
79. Kuhlman, P., Duff, H. L. & Galant, A. A fluorescence-based assay for multisubunit DNA-dependent RNA polymerases. *Anal. Biochem.* **324**, 183–190 (2004).
80. Kozlov, M., Bergendahl, V., Burgess, R., Goldfarb, A. & Mustaev, A. Homogeneous fluorescent assay for RNA polymerase. *Anal. Biochem.* **342**, 206–213 (2005).
81. Ishiura, M., Hazumi, N., Koide, T., Uchida, T. & Okada, Y. A recB recC sbcB recJ host prevents recA-independent deletions in recombinant cosmid DNA propagated in Escherichia coli. *J. Bacteriol.* **171**, 1068–1074 (1989).
82. Jones, I. M., Primrose, S. B. & Ehrlich, S. D. Recombination between short direct repeats in a recA host. *Mol. Gen. Genet. MGG* **188**, 486–489 (1982).

83. Azpiroz, M. F. & Laviña, M. Analysis of RecA-independent recombination events between short direct repeats related to a genomic island and to a plasmid in *Escherichia coli* K12. *PeerJ* **5**, e3293 (2017).
84. Bzymek, M. & Lovett, S. T. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8319–8325 (2001).
85. Fountain, F. F., Tolley, E. A., Jacobs, A. R. & Self, T. H. Rifampin hepatotoxicity associated with treatment of latent tuberculosis infection. *Am. J. Med. Sci.* **337**, 317–320 (2009).
86. Martínez, E., Collazos, J. & Mayo, J. Hypersensitivity reactions to rifampin. Pathogenetic mechanisms, clinical manifestations, management strategies, and review of the anaphylactic-like reactions. *Medicine (Baltimore)* **78**, 361–369 (1999).
87. Chen, J. & Raymond, K. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. *Ann. Clin. Microbiol. Antimicrob.* **5**, 3 (2006).
88. Ruslami, R. *et al.* Pharmacokinetics and Tolerability of a Higher Rifampin Dose versus the Standard Dose in Pulmonary Tuberculosis Patients. *Antimicrob. Agents Chemother.* **51**, 2546–2551 (2007).
89. Song, T. *et al.* Fitness costs of rifampicin resistance in *Mycobacterium tuberculosis* are amplified under conditions of nutrient starvation and compensated by mutation in the  $\beta'$  subunit of RNA polymerase. *Mol. Microbiol.* **91**, 1106–1119 (2014).
90. 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 (2011).
91. Banerjee, R., Rudra, P., Prajapati, R. K., Sengupta, S. & Mukhopadhyay, J. Optimization of recombinant *Mycobacterium tuberculosis* RNA polymerase expression and purification. *Tuberc. Edinb. Scotl.* **94**, 397–404 (2014).
92. Gregory, D., Tascon, R. E. & Lowrie, D. B. Repeated use of qiagen columns in large-scale preparation of plasmid DNA. *Methods Mol. Med.* **29**, 35–36 (2000).
93. Jacob, R. T. *et al.* MScreen: an integrated compound management and high-throughput screening data storage and analysis system. *J. Biomol. Screen.* **17**, 1080–1087 (2012).
94. Zhang, null, Chung, null & Oldenburg, null. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **4**, 67–73 (1999).
95. Murakami, K. S. X-ray crystal structure of *Escherichia coli* RNA polymerase  $\sigma 70$  holoenzyme. *J. Biol. Chem.* **288**, 9126–9134 (2013).

96. 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).
97. 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).
98. Römmele, G. *et al.* Resistance of Escherichia coli to rifampicin and sorangicin A—a comparison. *J. Antibiot. (Tokyo)* **43**, 88–91 (1990).
99. Xu, M., Zhou, Y. N., Goldstein, B. P. & Jin, D. J. Cross-resistance of Escherichia coli RNA polymerases conferring rifampin resistance to different antibiotics. *J. Bacteriol.* **187**, 2783–2792 (2005).
100. Oballa, R. M. *et al.* A generally applicable method for assessing the electrophilicity and reactivity of diverse nitrile-containing compounds. *Bioorg. Med. Chem. Lett.* **17**, 998–1002 (2007).
101. Garuti, L., Roberti, M. & Bottegoni, G. Irreversible protein kinase inhibitors. *Curr. Med. Chem.* **18**, 2981–2994 (2011).
102. Kathman, S. G., Xu, Z. & Statsyuk, A. V. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. *J. Med. Chem.* **57**, 4969–4974 (2014).
103. Young, R. M. & Staudt, L. M. Ibrutinib Treatment of CLL: The Cancer Fights Back. *Cancer Cell* **26**, 11–13 (2014).
104. Bauer, R. A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug Discov. Today* **20**, 1061–1073 (2015).
105. Emanuele, A. A. & Garcia, G. A. Mechanism of Action and Initial, In Vitro SAR of an Inhibitor of the Shigella flexneri Virulence Regulator VirF. *PLOS ONE* **10**, e0137410 (2015).
106. Cho, S., Lee, H. S. & Franzblau, S. Microplate Alamar Blue Assay (MABA) and Low Oxygen Recovery Assay (LORA) for Mycobacterium tuberculosis. *Methods Mol. Biol. Clifton NJ* **1285**, 281–292 (2015).
107. Falzari, K. *et al.* In vitro and in vivo activities of macrolide derivatives against Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **49**, 1447–1454 (2005).
108. Zhou, J. *et al.* Diminazene or berenil, a classic duplex minor groove binder, binds to G-quadruplexes with low nanomolar dissociation constants and the amidine groups are also critical for G-quadruplex binding. *Mol. Biosyst.* **10**, 2724–2734 (2014).
109. Baulard, A. R. *et al.* Activation of the pro-drug ethionamide is regulated in mycobacteria. *J. Biol. Chem.* **275**, 28326–28331 (2000).

110. Konno, K., Feldmann, F. M. & McDermott, W. Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am. Rev. Respir. Dis.* **95**, 461–469 (1967).
111. Metcalfe, C. *et al.* The tuberculosis prodrug isoniazid bound to activating peroxidases. *J. Biol. Chem.* **283**, 6193–6200 (2008).
112. Rybniker, J. *et al.* Lansoprazole is an antituberculous prodrug targeting cytochrome bc1. *Nat. Commun.* **6**, 7659 (2015).
113. Fleming, F. F., Yao, L., Ravikumar, P. C., Funk, L. & Shook, B. C. Nitrile-containing pharmaceuticals: efficacious roles of the nitrile pharmacophore. *J. Med. Chem.* **53**, 7902–7917 (2010).
114. Le Questel, J.-Y., Berthelot, M. & Laurence, C. Hydrogen-bond acceptor properties of nitriles: a combined crystallographic and ab initio theoretical investigation. *J. Phys. Org. Chem.* **13**, 347–358 (2000).
115. Janin, Y. L. Antituberculosis drugs: ten years of research. *Bioorg. Med. Chem.* **15**, 2479–2513 (2007).
116. Bastian, I. & Colebunders, R. Treatment and prevention of multidrug-resistant tuberculosis. *Drugs* **58**, 633–661 (1999).
117. Campbell, E. A. *et al.* Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* **104**, 901–912 (2001).
118. Barry, P. J. & O'Connor, T. M. Novel agents in the management of Mycobacterium tuberculosis disease. *Curr. Med. Chem.* **14**, 2000–2008 (2007).
119. Lounis, N. & Roscigno, G. In vitro and in vivo activities of new rifamycin derivatives against mycobacterial infections. *Curr. Pharm. Des.* **10**, 3229–3238 (2004).
120. Yamane, T. *et al.* Synthesis and biological activity of 3'-hydroxy-5'-aminobenzoxazinorifamycin derivatives. *Chem. Pharm. Bull. (Tokyo)* **41**, 148–155 (1993).
121. Walter, G., Zillig, W., Palm, P. & Fuchs, E. Initiation of DNA-dependent RNA synthesis and the effect of heparin on RNA polymerase. *Eur. J. Biochem.* **3**, 194–201 (1967).
122. Chen, Y.-J. *et al.* Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659–664 (2013).
123. Artsimovitch, I. *et al.* Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. *Cell* **122**, 351–363 (2005).



124. Nigam, A. *et al.* Modification of rifamycin polyketide backbone leads to improved drug activity against rifampicin-resistant *Mycobacterium tuberculosis*. *J. Biol. Chem.* **289**, 21142–21152 (2014).
125. Lin, W. *et al.* Structural Basis of *Mycobacterium tuberculosis* Transcription and Transcription Inhibition. *Mol. Cell* **66**, 169–179.e8 (2017).
126. Milstein, M. *et al.* Evaluation of high-dose rifampin in patients with new, smear-positive tuberculosis (HIRIF): study protocol for a randomized controlled trial. *BMC Infect. Dis.* **16**, (2016).
127. Burki, T. Multidrug resistant tuberculosis: a continuing crisis. *Lancet Infect. Dis.* **16**, 1337–1338 (2016).
128. Dooley, K. E., Nuermberger, E. L. & Diacon, A. H. Pipeline of drugs for related diseases: tuberculosis. *Curr. Opin. HIV AIDS* **8**, 579–585 (2013).
129. World Health Organization. *The stop TB strategy: building on and enhancing DOTS to meet the TB-related millennium development goals.* (World Health Organization, 2006).