Novel Molecular Mechanisms of Transcription Targeting Antibiotics

John Harbottle

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Institute for Cell and Molecular Biosciences

Faculty of Medicine

Newcastle University

January 2020

Abstract

Transcription, the first stage of gene expression, is performed by the multi-subunit RNA polymerase (RNAP). The indispensable nature of transcription and sequence divergence from eukaryotic counterparts make bacterial RNAP an excellent target for antibiotics. Yet very few clinical antibiotics target RNAP. The growing prevalence of antibiotic resistance amongst pathogenic bacteria demands the identification of novel antibacterial compounds acting through novel molecular mechanisms. This work consists of four distinct projects in which we investigated the molecular mechanisms of several previously uncharacterised transcription inhibitors.

(i)   Most clinical antibiotics are derived from the natural products of actinomycete bacteria. Consequently, we screened a library of actinomycetes compiled by our industrial collaborators Demuris™ for producers of novel inhibitors of bacterial transcription. From this screen we identified Antibiotic A39079S-1, produced by Streptomyces strain DEM40380, as an inhibitor of bacterial RNAP. The compound is an ansamycin type antibiotic with a previously uncharacterised mechanism of action. Here, we show the compound inhibits bacterial RNAP through a steric occlusion mechanism typical of rifamycins.

(ii)   Recently, the antibiotic ureidothiophene (Urd) was identified within a commercial screen of synthetic compounds in which inhibition of *S. aureus* RNAP was analysed. Here, we characterised the molecular mechanism of action by which Urd inhibits bacterial RNAP. We show the inhibitor targets regulatory sub-region 1.2 of the sigma subunit to prevent melting of the -10 promoter element. Consequently, Urd inhibits formation of the open promoter complex, a key step in transcription initiation.

(iii)   A prior screening program conducted by Demuris™ had identified the rifamycin type natural product kanglemycin A (KglA) as an inhibitor of rifampicin resistant RNAPs. Here, we show the unique ansa-bridge substituents of the compound act to form new binding contacts with RNAP. We also show

KglA inhibits transcription through a unique steric occlusion mechanism by preventing extension of the nascent transcript at an earlier stage than rifampicin.

(iv)   Finally, we investigated ADP-ribosylation as a mechanism of KglA inactivation by *Mycobacterium smegmatis* and *Mycobacterium abscessus* Rifampin ADP-ribosyltransferase (Arr) enzymes. We show KglA is not a substrate of the rifampicin inactivating Arr purified from *Mycobacterium smegmatis*, but remains a substrate of Arr purified from the extensively drug resistant pathogen *Mycobacterium abscessus*.

Declarations

a) I declare that this thesis is my own work and that I have correctly acknowledged the work of others. This submission is in accordance with University and School guidance on good academic conduct.

b) I certify that no part of the material offered has been previously submitted by me for a degree or other qualification in this or any other University.

c) I confirm that the word length is within the prescribed range as advised by my school and faculty

d) Does the thesis contain collaborative work, whether published or not?

Yes. Data obtained by others is clearly acknowledged in the thesis.

Signature of candidate .... Date;

Acknowledgements

List of Presentations and Publications

Presentation of this Work

- ❖ The Mechanism of Action of Ureidothiophene (Poster Presentation), 30th UK RNAP Conference, Charles Darwin House, The Biochemical Society, London (2018)
- ❖ Ureidothiophene Inhibits Open Complex Formation by Targeting Sigma Subunit Region 1.2 (Oral Presentation), 31st UK RNAP Conference, Clarendon Laboratory, Oxford University (2019)

Publication of this Work

- ❖ Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant Mycobacterium tuberculosis, Molecular Cell (2018).

  Hamed Mosaei*, Vadim Molodtsov*, Bernhard Kepplinger*, **John Harbottle**, Christopher William Moon, Rose Elizabeth Jeeves, Lucia Ceccaroni, Yeonoh Shin, Stephanie Morton-Laing, Emma Claire Louise Marrs, Corinne Wills, William Clegg, Yulia Yuzenkova, John David Perry, Joanna Bacon, Jeff Errington, Nicholas Edward Ellis Allenby, Michael John Hall, Katsuhiko S. Murakami and Nikolay Zenkin.

- ❖ Mechanisms of antibiotics inhibiting bacterial RNA polymerase, Biochemical Society Transactions (2019)

  Hamed Mosaei* & **John Harbottle***

- ❖ Ureidothiophene inhibits melting of -10 promotor element by targeting regulatory region 1.2 of sigma subunit, *submitting to Nucleic Acids Research (2020)*

  John Harbottle & Nikolay Zenkin

*equal contribution

Contents

List of Figures

List of Abbreviations

ATP - Adenosine Triphosphate

ADP-Rif – ADP-Ribosyl Rifampicin

ADP-Rifq – ADP-Ribosyl Rifampicin Quinone

ADP-KglA – ADP-Ribosyl Kanglemycin A

Arr – Rifampicin ADP-ribosyltransferase

Arr_{Ms} – *Mycobacterium smegmatis* Rifampicin ADP-ribosyltransferase

Arr_{Mab} – *Mycobacterium abscessus* Rifampicin ADP-ribosyltransferase

BH - Bridge Helix

$BH-H_N$ - Bridge Helix- N Terminal

Bp- Base Pair

Cap - Capistruin

Carb - Carbenicillin

Ci - Curie

Cor - Corallopyronin

CTP - Cytidine Triphosphate

DEM40380-CE – Crude cellular extract from *Streptomyces* DEM40380

DEM40380-Fn – Fraction of n% MeOH elution of C18 Solid Phase Extraction purification of crude cellular extract from *Streptomyces* DEM40380

DNA - Deoxyribonucleic Acid

DMSO - Dimethyl Sulfoxide

DTT - Dithiothreitol

EC - Elongation Complex

EDTA - Ethylenediaminetetraacetic acid

EMSA – Electrophoretic Mobility Shift Assay

ESI-MS Electro Spray Ionization-Mass Spectrometry

EtOH - Ethanol

Fdx - Fidaxomicin

Fnorm – Normalised fluorescence

GE - GE23077

GTP - Guanosine Triphosphate

GYM - Glucose-Yeast-Maltose

HPLC - High-performance liquid chromatography

IC_{50} - Half maximal inhibitory concentration

IPTG - Isopropyl β-D-1-thiogalactopyranoside

Kan – Kanamycin

KglA – Kanglemycin A

ADP-KglA – ADP-Ribosyl Kanglemycin A

L - Litres

LB – Luria-Bertani

LC-MS - Liquid Chromatography Mass Spectrometry

MccJ25 - Microcin J25

MeOH - Methanol

MIC - Minimum Inhibitory Concentration

Min - Minutes

MOA - Mechanism of Action

MRSA - Methicillin Resistant Staphylococcus aureus

NAC - Nucleotide Addition Cycle

NAD - Nicotinamide Adenine Dinucleotide

NMP - Nucleotide Monophosphate

NT – Non-Template

MST - Microscale Thermophoresis

Myx – Myxopyronin

M13ori – Origin of Replication of M13 bacteriophage

NTP - Ribonucleotide Triphosphate

OD_n - Optical Density at wavelength n

PCR - Polymerase Chain Reaction

PNK – T4 Polynucleotide Kinase

Polymin P - Polyethyleneimine Polymer

PUM - Pseudouridimycin

Rif – Rifampicin

Rif-S – Rifamycin S

Rif^R Rifampicin Resistant

Rifq – Rifampicin Quinone

Rip - Ripostatin

RNA - Ribonucleic Acid

RNAP - RNA polymerase

RPitc – Initially Transcribing Promoter Complex

RPo – Open Promoter Complex

RPc – Closed Promoter Complex

RP$_{i1}$ – Promoter intermediate complex 1

RP$_{i2}$ – Promoter intermediate complex 2

RRDR-N - Rifampicin Resistant Determining Region- N-terminus

RRDR-I - Rifampicin Resistant Determining Region I

RRDR-II - Rifampicin Resistant Determining Region II

RRDR-III - Rifampicin Resistant Determining Region III

rt – room temperature

S - Seconds

Sal - Salinamide A

SAR - Structure-Activity Relationship

PAGE - Polyacrylamide Gel Electrophoresis

SE - Standard Error

Sor – Sorangicin

SPE - Solid Phase Extraction

Stl – Streptolydigin

SQ – Squaramides

SW – Switch region

SW1 – Switch - 1

SW2 – Switch - 2

SW3 – Switch - 3

SW4 – Switch - 4

SW5 – Switch - 5

TB - Transcription Buffer

TGED - Tris-Glycerol-EDTA-DTT

TL - Trigger Loop

TLH1 – Trigger Loop Helix 1

TLH2 – Trigger Loop Helix 2

Tris - Tris(hydroxymethyl)aminomethane

TSS - Transcription Start Site

UTP - Uracil Triphosphate

Urd - Ureidothiophene

WT - Wild Type

X-gal - 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

σR – σ factor subregion

Chapter 1. Introduction

The first step of gene expression involves the conversion of genetic information encoded in a DNA template into a molecule of RNA. This process is called transcription. In all cellular organisms, transcription is performed by the multi-subunit enzyme DNA-dependent RNA polymerase (RNAP). RNAP catalyses the synthesis of RNA from a DNA template using ribonucleoside triphosphates. (NTPs) as substrates.

The essential nature of RNAP makes it an excellent target for therapeutic intervention. Indeed, inhibition of RNAP is an established approach in broad spectrum antibacterial therapy and antitubercular therapy (Villain-Guillot et al., 2007a, Darst, 2004, Chopra, 2007, Ma et al., 2016, Mosaei and Harbottle, 2019). Furthermore, understanding the mechanisms by which RNAP is targeted at a molecular level is of great value in the development of novel therapeutics. The expanding public health threat of antibiotic resistance demands the discovery of novel antibiotics with which to target resistant pathogenic bacteria. Additionally, understanding bacterial mechanisms of resistance is becoming increasingly important for effective drug development.

This introduction aims to briefly highlight the nature of transcription in bacteria and how RNAP is targeted by antimicrobial inhibitors. Most inhibitors covered in this work target initiation of transcription. Thus, there will be particular emphasis on processes of transcription initiation. Mechanisms of antibiotic resistance will also be briefly highlighted. Additionally, individual results chapters will also contain a brief introduction and discussion section.

1.1 The transcription cycle

The process of transcription in bacteria occurs in 3 distinct phases of i) initiation ii) elongation and iii) termination.

1.1.1 Initiation

During initiation of transcription in bacteria, RNAP is required to identify and bind specific duplex DNA sequences termed promoters. The 5 sub-unit catalytic core

RNAP (α2β'βω) is unable to achieve this independently (Ruff et al., 2015, Burgess et al., 1969). The core enzyme requires an additional initiation factor for recruitment to promoter sequences and the melting of duplex DNA. In bacteria, this accessory factor is the σ factor which binds to the core enzyme to form the holoenzyme (Burgess and Anthony, 2001, Murakami and Darst, 2003, Borukhov and Nudler, 2003, Burgess et al., 1969). The primary σ factor (σ^{70} in *Escherichia coli*) directs transcription to the majority of promoters during exponential growth phase (Paget and Helmann, 2003). Alternative σ factors control transcription of genes in response to certain environmental conditions (Paget, 2015).

Following binding of σ to core, σ directs the resultant holoenzyme to a set of promoters dictated by the sequence specificity of the given σ factor (Figure 1.1). The σ^{70} holoenzyme recognises cognate -10 (consensus: 5'-TATAAT-3') and -35 (consensus: 5'-TTGACA-3') promoter hexamers upstream of the transcription start site (TSS, denoted as +1) (Paget and Helmann, 2003, Murakami, 2013, Burgess and Anthony, 2001). σ^{70} is comprised of five distinct modular domains (σR1.1, σR2, σR3, σR3.2 and σR4) of which σR2 and σR4 are responsible for the recognition of the cognate -10 and -35 promoter elements, respectively (Campbell et al., 2002a). Binding of duplex promoter DNA to RNAP constitutes the formation of the closed promoter complex (RPc) in which promoter DNA remains double stranded (Ruff et al., 2015, Mazumder and Kapanidis, 2019, Hawley and McClure, 1980, Li and McClure, 1998). Following formation of RPc, σ^{70} nucleates promoter melting by intercalating aromatic residues into DNA at the -12 position of duplex DNA (Feklistov and Darst, 2009). Conserved non-template DNA bases (NT) at the −11 and −7 positions are flipped out of the DNA duplex and stabilised in protein pockets on σR2 (Zhang et al., 2012, Murakami et al., 2002b, Bae et al., 2015a). The template strand is loaded into the main cleft of RNAP to position the TSS within the active centre. The melted portion of DNA, known as the transcription bubble, extends from the -11 position to the +2 position (Murakami, 2015, Bae et al., 2015a). RNAP complexed with melted promoter DNA is termed the open promoter complex (RPo). RNAP is then competent for binding of NTPs and the synthesis of RNA transcripts.

Before transitioning into processive elongation, the open complex undergoes a period of abortive transcription in which the RNAP retains promoter specific

contacts and reiteratively synthesises short RNAs, this species is termed the initially transcribing complex (RPitc) (Figure 1.1). This reiterative abortive synthesis is achieved by a DNA scrunching mechanism, in which DNA is pulled into the active centre whilst the trailing edge of the enzyme remains static on the DNA template (Kapanidis et al., 2006). Energy stored in this 'scrunched complex' is resolved either by the release of the short nascent RNA transcript as an abortive product and the return of the enzyme to the open promoter complex, or the energy of the scrunched complex causes RNAP to relinquish its contacts with the promoter, dissociate from the σ factor and transition to a highly stable elongation complex (EC) (Gralla et al., 1980, Carpousis and Gralla, 1980). (See section 1.2 for further detail regarding initiation of transcription)
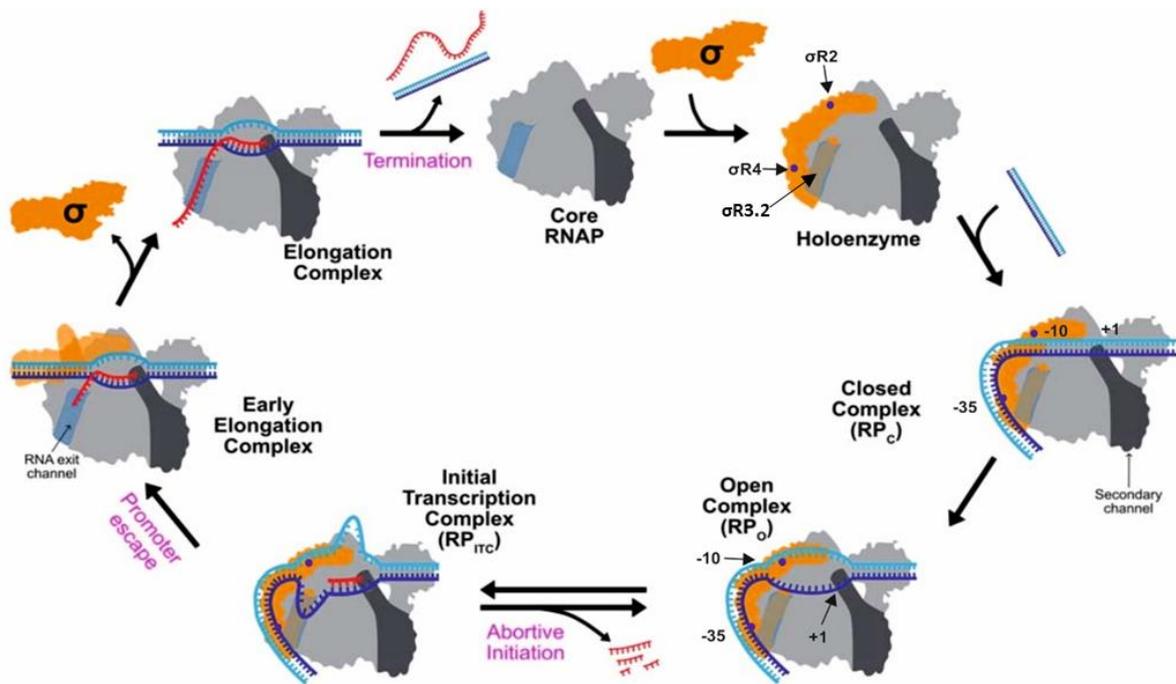
Figure 1.1 Schematic of the transcription cycle and σ:DNA interactions. Adapted from (Alhadid et al., 2017). The beginning of the transcription cycle is delineated by the binding of σ (orange) to core RNAP (grey) to form the RNAP holoenzyme. The secondary channel (dark grey) and RNA exit channel (light grey) are depicted on the right and left hand sides of RNAP, respectively. σR3.2 is depicted within the RNA exit channel. Holoenzyme binds double stranded promoter DNA to form the closed promoter complex (RPc) by recognition of the -10 and -35 promoter elements by σR2 and σR4 (purple dots), respectively. RPc undergoes spontaneous isomerisation to the transcription competent open promoter complex (RPo). RPo then undergoes a period of reiterative abortive synthesis in which short RNAs are synthesised and released (RPitc). Eventually, RNAP escapes the promoter by displacing σR3.2, relinquishing promoter specific contacts, and dissociating from sigma, before transition to a highly processive elongation complex. (The author directs the reader to section 1.2 for further details regarding the transcription cycle)

1.1.2 Elongation

During elongation, the EC travels along the DNA template synthesising a complementary RNA transcript. Transcription elongation in prokaryotes is extraordinarily processive; the enzyme is capable of synthesising transcripts thousands of nucleotides in length without disassembly of the EC (Nudler et al., 1996). Extension of RNA occurs through the nucleotide addition cycle (NAC) (Figure 1.2). NAC is marked by binding of the NTP substrate within the active centre followed by nucleotidyl transfer at the growing 3' end of the nascent transcript. Thus, the RNAP active centre contains a binding site for the 3' terminus of RNA (i site) and a binding site for the incoming NTP (i+1 site) (Vassylyev et al., 2007a, Korzheva et al., 2000). Following nucleotide addition, release of pyrophosphate generates an EC in which the 3' end of RNA is bound in the i+1 site of the active centre. This is termed the pre-translocated state. To ensure further elongation of the transcript, the EC must translocate, positioning the 3' end of the transcript into the i site and vacate the i+1 site for an incoming NTP substrate. This is termed the post-translocated state (Korzheva et al., 2000)

Throughout elongation, the EC transiently maintains a transcription bubble of ~12 bp containing a 9-10bp RNA-DNA hybrid (Nudler et al., 1997). Translocation is accompanied by 1bp melting of downstream duplex DNA and a corresponding 1bp restoration of duplex DNA upstream. Additionally, translocation of RNAP causes the RNA-DNA hybrid to shorten by 1bp. This shortening of the RNA-DNA hybrid, combined with movement of the transcription bubble, makes passive translocation energetically unfavourable (Nudler, 2009, Bar-Nahum et al., 2005a). Therefore, it is thought NTP substrate plays a role in assisting RNAP translocation. Within this model, RNAP oscillates between pre- and post-translocated states with the 3' end of RNA moving between the i and i+1 site. Binding of NTP stabilises the post translocated state by occupying the vacant site and locking the enzyme in a catalytically competent state (Bar-Nahum et al., 2005b). The substrate acts as a stationary pawl in the 'Brownian ratchet' mechanism of RNAP translocation. Subsequent nucleotidyl transfer and pyrophosphate release act to reduce the thermodynamic stability of the EC owing to the significant binding energy afforded by the pyrophosphate moiety in NTP

binding. Consequently, upon pyrophosphate release, the EC can move laterally, adopt the post-translocated state and repeat the nucleotide addition process.
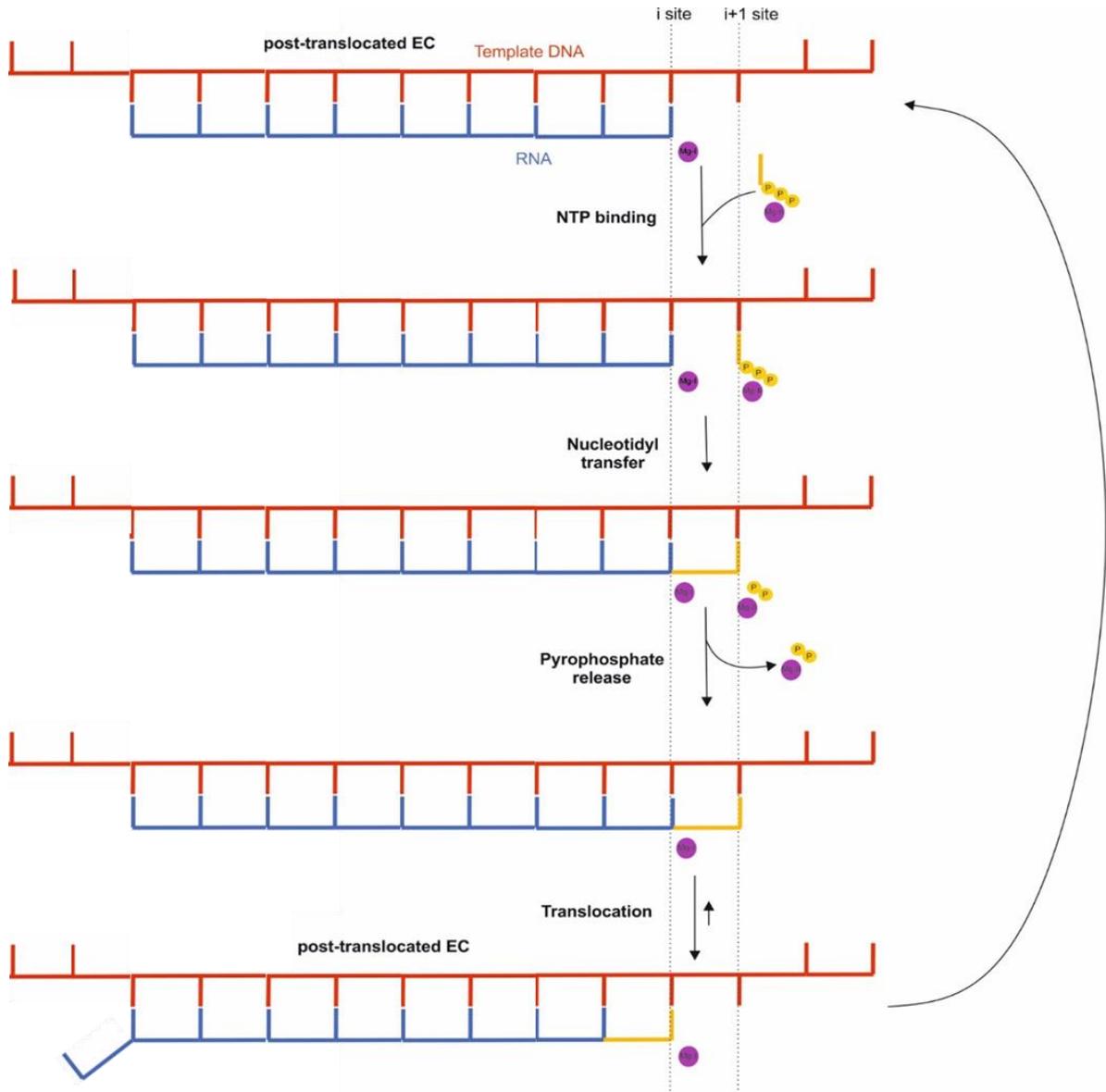
Figure 1.2 The nucleotide addition cycle (NAC). Template DNA is shown in red, RNA is shown in blue, incoming nucleotides are shown in yellow, and catalytic Mg^{2+} shown in magenta. Non-template is omitted for clarity. At the beginning of the NAC, RNAP is in the post-translocated state in which RNA 3'-OH is located in the active centre i site. NTPs diffuse into the active site with chelated Mg-II, and bind the i+1 site. RNAP then catalyses phosphodiester bond formation, following which a pyrophosphate molecule is released. Following catalysis, RNAP then translocates, vacating the i+1 site and manoeuvring the RNA 3'-OH into the i site in preparation for the subsequent nucleotide addition cycle.

Principally, polymerisation of RNA occurs through the nucleophilic attack of the 3'
hydroxyl group of the RNA strand on the α-phosphate of an incoming NTP. This
S_N2-type reaction results in the formation a new phosphodiester bond and
displacement of pyrophosphate. The RNAP active centre, like all nucleic acid
polymerases, catalyses this process through a two metal ion mechanism
(Sosunov et al., 2003) (Figure 1.4). RNAP co-ordinates a catalytic Mg^{2+} ion (Mg-I)
by an evolutionarily conserved invariant aspartate triad of β'. Principally, Mg-I is
tasked with activation of the 3' terminus hydroxyl group and correct positioning of
the incoming NTP through coordination with the α-phosphate. A second Mg^{2+}
(Mg-II) is bound by RNAP ~ 6Å away from Mg-I and with lower affinity. It is
thought to be delivered to the active centre bound to incoming NTP, stabilised by
the α-, β- and γ- phosphates. The role of Mg-II is to coordinate the negatively
charged oxygens of the α-, β- and γ- phosphates, ensuring their correct
orientation for the nucleophilic attack of the 3' hydroxyl group (Nudler, 2009,
Sosunov et al., 2003). The detailed structure of the active centre and its role in
catalysis will be discussed further in 1.2.1.

1.1.3 Termination

Transcription termination in prokaryotes occurs through 2 distinct mechanisms;
intrinsic termination (Rho-independent) and Rho-dependent termination. In
intrinsic termination, RNAP transcribes a palindromic G-C rich region followed by
a U-rich tract. Following transcription of the U-rich tract, RNAP pauses. This
pause allows the G-C rich region of the transcript to form a hairpin structure
within the RNA exit channel. The hairpin extends into the main channel causing a
shortening of the RNA-DNA hybrid and displacing RNA from the exit channel.
The destabilised EC then dissociates into its constituent parts for further rounds
of transcription (Gusarov and Nudler, 1999).

In Rho-dependent termination, the homohexameric Rec-A family helicase Rho
binds preferentially to C-rich regions of transcribed RNA. RNA is threaded
through Rho's central pore, triggering 5' to 3' RNA translocase activity (Pallabi et
al., 2017, Ananya et al., 2016). Rho translocates towards RNAP halted at Rho
dependent termination sites. Rho then interacts with RNAP to inactivate and
destabilise the stationary EC causing it to dissociate (Pallabi et al., 2017).

1.2 Structure & Mechanisms of RNAP

1.2.1 Overall core RNAP structure

The core RNAP is composed of 5 subunits (α2ββ'ω) with an overall molecular weight of ~ 400KDa (Murakami, 2015, Korzheva et al., 2000, Vassylyev et al., 2002, Borukhov and Nudler, 2003, Zhang et al., 1999). The enzyme possesses a distinctive crab claw structure with the two largest subunits β and β' constituting the opposing pincer-like structures (Figure 1.3, left panel). The β and β' subunits delineate the main channel of RNAP with the active centre located towards the rear of the claw. The two α subunits are located at the back of the enzyme. They form a homodimer through interactions of their N-termini onto which the β and β' subunits are loaded during RNAP assembly (Ishihama, 1992). The ω subunit also plays a role in assembly, it is thought to aid in recruitment of the β' subunit during assembly and stabilise its conformation within the fully assembled RNAP (Mathew and Chatterji, 2006).
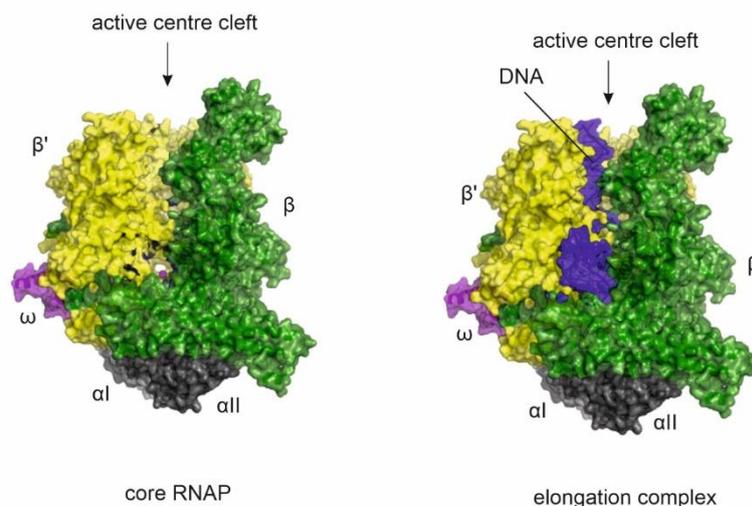
Figure 1.3 Overall structure of *E. coli* core RNAP and the RNAP elongation complex (surface representations). (Left) The structure of the RNAP core enzyme. β is depicted in yellow, β' is depicted in green, ω is depicted in purple, α-II is depicted in dark grey, and α-I depicted in light grey. (Right) The structure of the *E. coli* RNAP elongation complex. Individual subunits are depicted as in core RNAP, with DNA depicted in blue. (Images constructed from PDB file 6RH3)

Duplex DNA enters RNAP through the main channel formed by the β and β' subunits (Figure 1.3, right panel). As the DNA nears the active centre it is melted

into the transcription bubble and the template strand is passed through the active centre to direct complimentary RNA synthesis (Vassylyev et al., 2007a, Korzheva et al., 2000). During transcription, the main channel houses a transiently maintained 9-10bp RNA-DNA hybrid (Mustaev et al., 2017, Nudler et al., 1997). Once RNA extends beyond 10bp, the 5' end is threaded through the RNA exit channel. Access of NTP substrate to the active centre is thought to be enabled through the secondary channel. Also termed the pore, the secondary channel is a ~12Å wide channel that extends from the active centre to the surface of the enzyme. NTPs are thought to diffuse through the secondary channel to the active site where they are selected for incorporation at the 3' end of RNA (Korzheva et al., 2000, Batada et al., 2004). Additionally, the secondary channel accommodates the 3' end of RNA during backtracking of RNAP (Mustaev et al., 2017, Korzheva et al., 1998). However, there is also some evidence to suggest that NTPs are able to access the active site through the main channel (Landick, 2005, Burton et al., 2005).

Within the overall structure, RNAP possesses a number of mobile elements vital to the enzyme function. The majority of the β' subunit constitutes the larger of the two pincers, termed the 'clamp' (Chakraborty et al., 2012). It is a highly mobile domain able to hinge around a flexible region at its base termed the switch region comprised of five discrete elements (SW1–SW5) (Belogurov et al., 2009, Mukhopadhyay et al., 2008). The clamp is able to swing open to accommodate the DNA within the active centre cleft. Upon binding of DNA, the clamp reverts to a closed conformation to retain melted DNA within the active centre during initiation and elongation (Chakraborty et al., 2012, Duchi et al., 2018, Feklistov et al., 2017). Within the main channel there are a number of other structural motifs that influence the transcription process. The β flap domain, β' lid, β' zipper, and Zn^{2+} finger domain contribute to the structure of the RNA exit channel. These domains contribute to the stability and length of the RNA-DNA hybrid within the main channel and displace RNA from the DNA template during elongation (Korzheva and Mustaev, 2001, Vassylyev et al., 2007a). The β' lid functions as an upstream zip lock serving to displace RNA and prevent overextended hybridisation with the template-strand. The β' rudder domain further stabilises the elongation complex by interacting with the upstream edge of the hybrid. A key

structural element of the active centre termed the bridge helix (BH) acts along with the β fork loop-II to serve as the downstream zip lock (see section 1.2.2 for further details on BH function). Following de-hybridsation at the upstream edge, the RNA is channelled into a narrow cavity formed by the β' lid, β' Zn^{2+} finger and the β flap domain. The β lobes accommodate the non-template strand during elongation (Korzheva et al., 2000).

1.2.2 Architecture of the active centre

The active site of RNAP, denoted by catalytic Mg-I and Mg-II ions, is located within the main cleft at the rear of the crab claw (Vassylyev et al., 2002, Vassylyev et al., 2007b). As shown in Figure 1.4, Mg-I is tasked with activation of the 3' RNA hydroxyl group and positioning the incoming NTP through coordination with the α-phosphate. Whereas Mg-II is charged with coordinating the α-, β- and γ- phosphates, ensuring their correct orientation for the nucleophilic attack of the 3' hydroxyl group (Sosunov et al., 2003) . In doing so, Mg-I and Mg-II stabilise the pentavalent transitions state that typifies such S_N2-type reactions.

Mg-I is bound permanently within the active centre, coordinated through the aspartate triad of the universally conserved NADFDGD motif present in the β' subunit. Any substitution of these aspartate residues (β'D460, β'D462 and β'D464, *E. coli numbering*) abolishes all catalytic activity of RNAP (Sosunov et al., 2005, Zaychikov et al., 1996). Mg-II is thought to be brought to the active centre bound to the incoming NTP substrate and coordinated primarily by two of the catalytic aspartates (β'D460, β'D462). Consequently, Mg-II is bound to RNAP with significantly lower affinity than Mg-I (Sosunov et al., 2003).
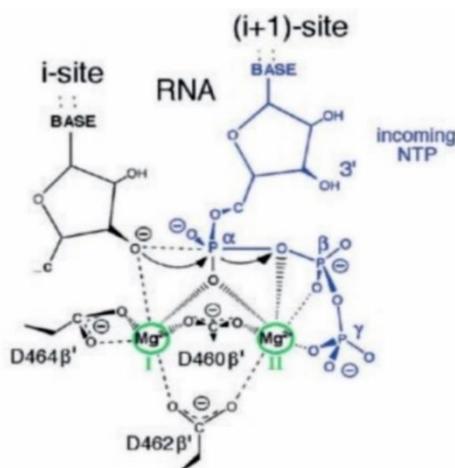
Figure 1.4 Catalytic mechanism of phosphodiester bond formation. Adapted from (Sosunov et al., 2005) Two catalytic magnesium ions (green) are coordinated by three β' aspartate residues (*E. coli* numbering). The incoming NTP substrate is in blue. Curly arrows depict the movement of electron pairs during the formation of the phosphodiester bond. Activated RNA 3'-hydroxyl group attacks the α- phosphate of the NTP bound in the i+1 site, forming a new phosphodiester bond, and displacing a pyrophosphate molecule.

Two structural domains of the active centre, the trigger loop (TL) and BH, are known to play crucial roles in catalysis in all RNAPs. The TL is a highly flexible domain located in the vicinity of the active centre that moves between closed and open conformations during catalysis (Mejia et al., 2015, Temiakov et al., 2005a, Zhang et al., 1999) (Figure 1.5). It, along with the BH, contributes to binding and selection of cognate NTP substrates (Vassylyev et al., 2007a, Vassylyev et al., 2007b, Wang et al., 2006). Firstly, the open TL conformation permits NTP entry into a 'pre-insertion site'. If bound NTP matches the base of the i+1 site, the TL refolds to the closed conformation. This closed state is stabilised through interactions of the TL with the base of cognate NTP substrate (Vassylyev et al., 2007b, Wang et al., 2006, Yuzenkova et al., 2010) . Incorrect hydrogen bonding of non-cognate substrates with the i+1 template base prevents complete TL refolding due to steric clash with the base of the incoming NTP (Wang et al., 2006, Vassylyev et al., 2007b, Yuzenkova et al., 2010). As a fully closed TL is a prerequisite for catalysis, the TL contributes to transcription fidelity. Furthermore, the TL directly contributes to catalysis through β' residues R933 and H936 (*E. coli* numbering). In the closed conformation, these residues are brought into close proximity to the catalytic Mg^{2+} ions and coordinate the α-, β- and γ- phosphates of the incoming NTP substrate, further stabilising the transition state during nucleotidyl transfer (Vassylyev et al., 2007b, Yuzenkova et al., 2010).
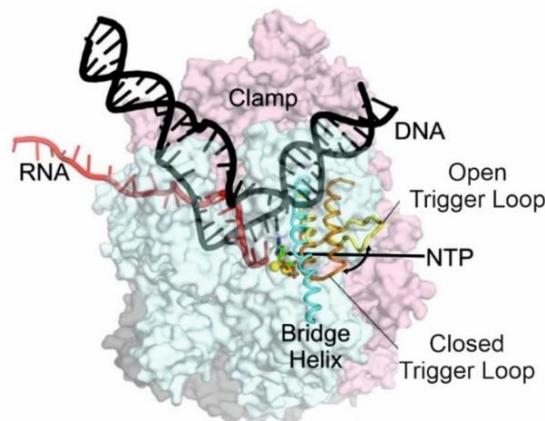
Figure 1.5 Location of the bridge helix (BH) and trigger loop (TL). Adapted from (Hein and Landick, 2010) Structure of the *T. thermophilus* elongation complex. DNA is depicted in black, RNA is depicted in red. RNA polymerase subunits are shown as partially transparent surface representations (α, β, and β'; blue, gray and pink, respectively). The positions of the BH (cyan) and closed (folded) TL (orange) are shown. Conformational change to the open (unfolded) TL (yellow) is indicated by the black double ended arrow. The thick black arrow indicates proposed NTP access through the secondary channel. Active site Mg^{2+} is indicated in yellow and α,β-methylene-ATP substrate indicated in green.

Alongside the TL, the BH plays a concerted role in loading of cognate NTP and the translocation of RNAP following phosphodiester bond formation. The BH is a large metastable α-helix which spans the main cleft of RNAP, bisecting the main cleft into the primary and secondary channels (Vassylyev et al., 2007a, Zhang et al., 1999, Korzheva et al., 2000). During catalysis, the BH is thought to transition between straight and kinked conformations. The TL and BH are proposed to work in tandem to enable translocation by a two pawl ratchet mechanism (Bar-Nahum et al., 2005a). Transition of the TL to the folded conformation is thought to induce a kink of the BH which may push the 3' nucleotide into the i site, vacating the i+1 site for further nucleotide addition. Consequently, the RNAP moves from the pre-translocated state to the post-translocated state. Additionally, the BH works alongside the TL to correctly position the incoming NTP for catalysis (Korzheva et al., 2000, Vassylyev et al., 2007a).

1.2.3 σ⁷⁰ and the holoenzyme

For sequence-dependent transcription initiation, the core enzyme must bind the initiation factor, σ (Figure 1.6). In recent years, a number of high resolution structures have described the exquisite architecture of the σ factor and its organisation within the bacterial holoenzyme (Murakami et al., 2002a, Murakami et al., 2002b, Murakami, 2013).

In the case of *E. coli,* σ⁷⁰ possesses a number of distinct modular domains; σR1.1, σR2, σR3, σR3.2 and σR4. σR2, σR3 and σR4 are structured globular domains adjoined by flexible linkers (Figure 1.7). σR3.2 constitutes a functionalised linker between σR3 and σR4, whereas σR1.1 is a negatively charged disordered domain (Campbell et al., 2002b). σR2, σR3 and σR4 constitute the binding interface between core and σ; σR2 interacts with β' clamp

domain both on the exterior and within the main channel, while σR3 and σR4 interact with the β flap domain. σR3.2 serves as a flexible linker domain situated between the globular domains of σR3 and σR4. It extends into the main channel and interacts with template DNA upstream of the active centre in the open promoter complex (Murakami, 2013, Murakami et al., 2002a, Murakami et al., 2002b). It is implicated in a number of aspects of transcription initiation (see below). σR1.1 is located within the main RNAP cleft within the holoenzyme, acting as a 'gatekeeper' by mimicking negatively charged DNA. Within the open promoter complex, promoter DNA displaces σR1.1 from the main channel.
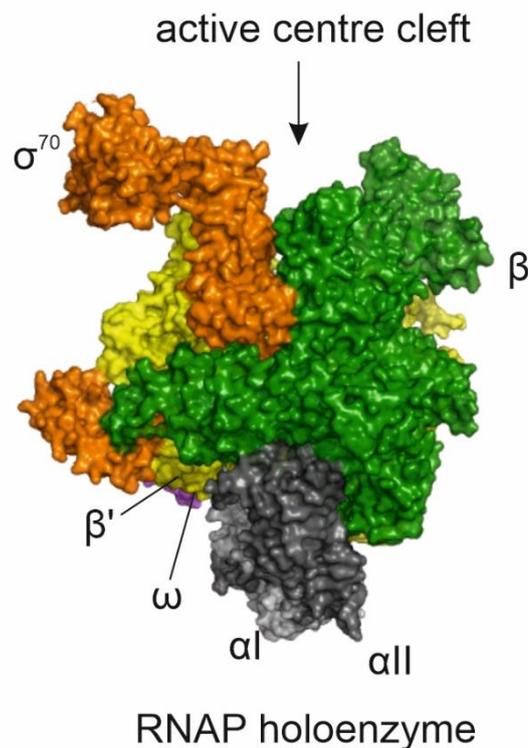
Figure 1.6 Overall structure of the *E. coli* holoenzyme (surface representation). β is depicted in yellow, β' is depicted in green, ω is depicted in purple, α-II is depicted in dark grey, α-I depicted in light grey, and σ70 depicted in orange (Images constructed from PDB file 6CA0).

1.2.4 Promoter architecture and the open complex (RPo)

During σ^{70}-dependent initiation, within the context of the holoenzyme, σR2 and σR4 engage with the -10 (consensus: 5'-TATAAT-3') and -35 (consensus: 5'-TTGACA-3') hexamers, respectively (Ruff et al., 2015, Zhang et al., 2012, Bae et al., 2015a, Feng et al., 2016, Dickson et al., 1975) (Figure 1.7). Although, note, in nature, almost all promoters diverge somewhat from these canonical sequences. A consensus spacer region of 17bps separates the -10 and 35 sequences, although this distance can vary from between 16-19bp, depending upon the promoter (Shimada et al., 2014, Hawley and McClure, 1983). Furthermore, the α-C terminal domains (αCTD) interacts with DNA upstream of the -35, at positions -40 to -60, termed the UP element (Ross et al., 1993, Gourse et al., 2000). This additional promoter element has been shown to influence rates of promoter complex formation and overall rates of transcription. Additionally, some promoters, such as the galP1 promoter, possess an additional motif (consensus: 5'-TG-3') 2 bps upstream of the -10 element which interacts with σR3 to further stabilise open promoter complexes (Barne et al., 1997, Haugen et al., 2006). Furthermore, the region between the TSS and -10, termed the discriminator region, is also thought to play a role in RPo lifetime (Haugen et al., 2006, Shimada et al., 2014).

As shown in Figure 1.7, within RPo, σR2 binds to the -10 element, σR3 binds to the extended -10 motif and σR4 binds to the -35 promoter element. In particular, a helix-turn-helix substructure of σR4, σR4.2, binds to bases of the -35 element (Zhang et al., 2012, Murakami et al., 2002b, Murakami et al., 2002a). A substructure of σR2, σR2.3, makes specific contacts with the conserved NT adenosine and thymine bases at the -11 position and -7 position, respectively (Zhang et al., 2012, Murakami et al., 2002b, Murakami et al., 2002a, Feng et al., 2016). The bases are flipped out of the duplex into protein pockets, while σ^{70} makes further non-specific contacts with other bases of the -10 elements to stabilise the open promoter complex.
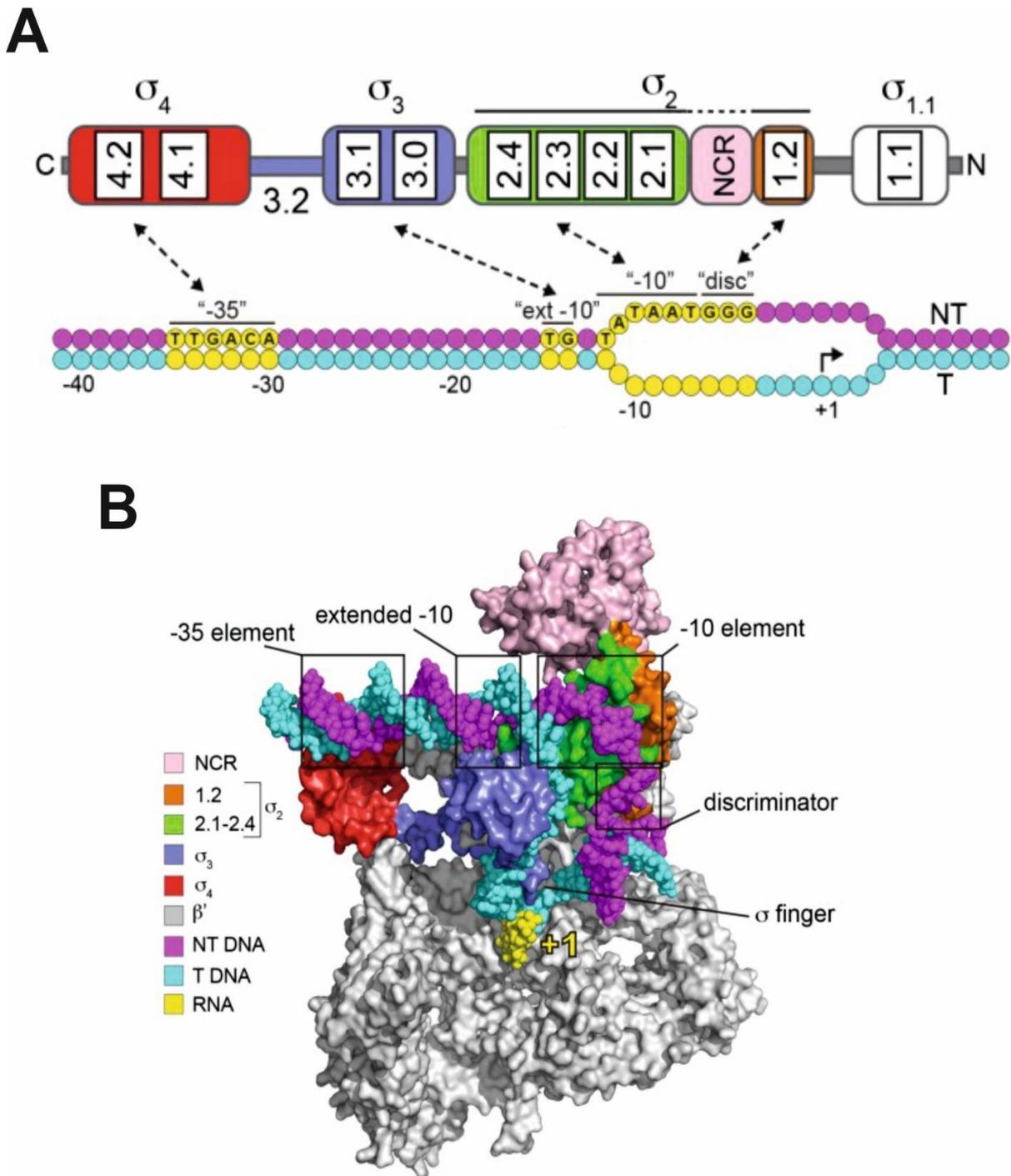
Figure 1.7 Domain architecture of σ70 and its interaction with promoter DNA within the open promoter complex. Adapted from (Paget, 2015). A) Schematic represents the overall globular domain structure of σ70 with respective subdomains noted and their interactions with promoter DNA in the context of the open promoter complex denoted by dashed arrows (RPo). σR4 (σ4) is denoted in red, σR3 (σ3) is denoted in dark blue, σR2 (σ2) is denoted in green, and σR1.2 is denoted in orange (NCR; non conserved region). In the schematic of promoter DNA below, individual consensus promoter elements are denoted relative to the transcription bubble and

transcription start site. The non-template (NT) strand is depicted in purple, and the template (T) strand is depicted in cyan. B) Structure of the *E. coli* transcription initiation complex. σ70 domains (surface representations) and DNA (sphere representations) are coloured as in A. For clarity the β, α-II and ω subunits are omitted.

A further substructure of σR2, σR1.2, interacts with non-template promoter DNA downstream of the -10 element within the discriminator region (Zhang et al., 2012, Haugen et al., 2006, Haugen et al., 2008a). The base at non-template strand position -6 is stabilised in a protein pocket on σR1.2 within the open complex, whilst interactions are also made with the -5 and -4 positions. Interestingly, σR1.2 facilitates the binding of single stranded -10 promoter DNA by σR2 through an allosteric signal within the context of the holoenzyme (Zenkin et al., 2007). Holoenzymes lacking σR1.2 are unable to recognise single stranded -10 promoter sequence DNA (Zenkin et al., 2007, Wilson and Dombroski, 1997). Furthermore, certain substitutions in σR1.2 make the holoenzyme incapable of forming stable open complexes and are highly defective in transcription initiation (Baldwin and Dombroski, 2001). It is hypothesised σR1.2 stabilises a particular σ factor conformation required for optimal -10 promoter element recognition (Zenkin et al., 2007).

1.2.5 Mechanism of open complex formation

Several promoter complex structures, inhibitor-promoter complex structures, single molecule experiments and footprinting data have allowed the elucidation of several intermediates that lie on the pathway towards RPo (Ruff et al., 2015, Roe et al., 1984, Buc and McClure, 1985, Mazumder and Kapanidis, 2019). Although individual promoters likely have varying intermediates and kinetics, there is a general consensus on how RPc transitions to RPo, and a simple sequence of events can generally be attributed to all promoters. Firstly, the holoenzyme must recognise promoter DNA to form an initial closed promoter complex intermediate, termed RP_{i1}, in which promoter DNA is double stranded (Figure 1.8, A). Subsequently, a rate-limiting isomerisation to a heparin resistant intermediate, termed I_2, occurs in which an unstable melted DNA bubble is formed, before a final transition to stable transcription competent RPo (Ruff et al., 2015, Bae et al., 2015a, Mazumder and Kapanidis, 2019a) (Figure 1.8, A,B).
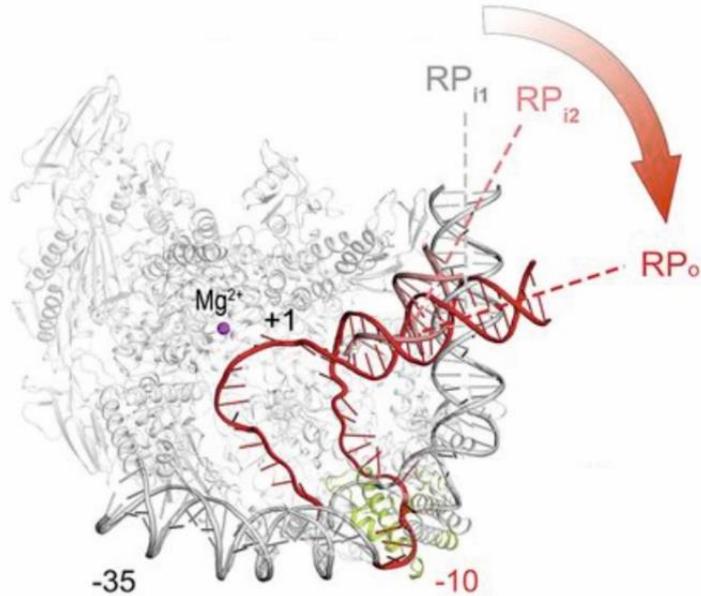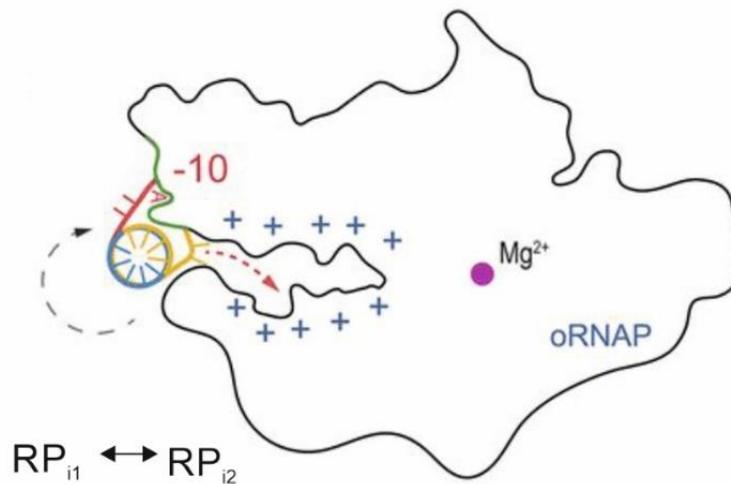
Figure 1.8 Formation of the open promoter complex (RPo). Adapted from (Feklistov et al., 2017) A) Kinetic scheme of promoter melting by RNAP (R, RNAP; P, promoter DNA; RP_{i1}, promoter melting intermediate 1; RP_{i2}, promoter melting intermediate 2, and RPo, open promoter complex. B) Modelled positions of downstream duplex DNA within intermediates of promoter melting. Respective intermediates and DNA trajectories are indicated with respect to the above kinetic equation. Promoter elements are indicated, σR2.3 is depicted in green. C) Schematic of spontaneous DNA unwinding during promoter melting. Blue symbols indicate positively charged interior surface of the active site cleft. σR2.3 is depicted in green, template DNA is depicted in yellow, non-template DNA is depicted in blue with -10 promoter element and conserved -11 adenosine depicted in red, and catalytic Mg^{2+} depicted in magenta. Schematic model of RP_{i1} to RP_{i2} transition is shown.

The slow isomerisation of $I_1 > I_2$ is thought to be driven by both thermal dynamics and an active role of RNAP. It is widely accepted the first step of promoter melting involves flipping-out of the conserved non-template -11 adenosine base into protein pockets on σR2 (Feng et al., 2016, Zhang et al., 2012, Bae et al., 2015a). This is thought to activate recognition of the remaining bases of the -10 element and allow promoter melting to propagate to the TSS. Several conserved aromatic residues of σR2.3, namely F419, Y430, W433 and W434 stabilise the flipped out base and act to bend promoter DNA 90° towards the RNAP active site cleft during melting (Mazumder and Kapanidis, 2019a, Feklistov and Darst, 2009, Koo et al., 2009). Y430 and W433 are implicated in intercalation of double stranded promoter DNA at the -12 position to initiate nucleation and stabilise the flipped out conformation of the -11 base (Feklistov and Darst, 2011).

Formation of RPo varies greatly depending upon both urea and salt concentration. This dependency suggests formation of stable RPo involves significant conformational changes of RNAP, namely the mobile β' clamp domain (Mekler et al., 2014, Drennan et al., 2012). Within RPo, the clamp adopts a 'closed' conformation, in which the entry of single stranded DNA is prohibited due to the width of the cleft and the interactions of σR2 with promoter DNA above the cleft (Feklistov et al., 2017, Boyaci et al., 2019). This indicates RPo formation requires dynamic movement of the clamp during promoter melting. Single molecule FRET studies have shown that the clamp can adopt several conformations; a 'closed', 'partly closed' and 'open' conformation, with the 'open' conformation predominating in solution (Chakraborty et al., 2012). Indeed, several current models suggest promoter melting is intricately linked with dynamics of the clamp (Boyaci et al., 2019, Feklistov et al., 2017, Lin et al., 2017a). It is now proposed, following recognition of upstream promoter elements, a transient, thermally driven closure of the clamp allows σR2 to engage with the -10 promoter element, consequently nucleating promoter melting above the cleft (Feklistov et al., 2017). As melting propagates downstream of the -10 towards the TSS, opening of the clamp (oRNAP) allows negatively charged single stranded template DNA to be pulled towards the active centre by electrostatic interactions with positively charged residues within the cleft (Feklistov et al., 2017) (Figure 1.8,C). This particular intermediate in which the DNA is melted, yet the clamp

remains open, likely represents the unstable open promoter intermediate I_2. Subsequent closure of the clamp stabilises the complex to form transcription competent RPo in which template DNA is bound stably within the active site.

1.2.6 Mechanism of initial transcription and promoter escape

In the presence of NTPs, an open promoter complex undergoes reiterative RNA synthesis which can result in either short abortive RNA synthesis or productive RNA synthesis (Kapanidis et al., 2006). The productive pathway is characterised by synthesis of an RNA molecule of 9-11nt in length, at which point the RNAP relinquishes contacts with the promoter and enters processive elongation. In contrast, the abortive pathway involves RNAP retaining contacts with the promoter, synthesising and releasing short RNAs (termed abortive products), returning to RPo, and then reinitiating synthesis of RNA (Carpousis and Gralla, 1980, Gralla et al., 1980). The balance between these two pathways is influenced by promoter sequence and the initially transcribed sequence (Hsu et al., 2003, Hsu, 2009). Indeed, it is thought to be an influential regulatory mechanism at certain promoters *in vivo.*

Utilisation of single molecule methods enabled identification of the mechanism by which RNAP proceeds through abortive transcription. By measuring distances within promoter complexes engaged in abortive RNA synthesis, it was discovered downstream DNA is pulled into RNAP 1bp per nucleotide addition cycle causing an enlargement of the transcription bubble, whilst the enzyme remains stationary on the promoter fragment (Revyakin et al., 2006). This so-called 'DNA scrunching' mechanism is an essential step in promoter escape. It is believed, during abortive transcription, accumulated DNA unwinding and compaction create a high energy intermediate in which stress accumulation drives the breakage of specific interactions between RNAP, promoter DNA and σ^{70}. Conversely, this high energy intermediate can be resolved by releasing the short RNA product and returning to RPo to reinitiate RNA synthesis.

Structural studies of the RNAP holoenzyme and initially transcribing complexes show the σ^{70} substructure σR3.2 is located in the path of extending RNA during initiation (Basu et al., 2014, Murakami et al., 2002a, Murakami et al., 2002b). Once a nascent transcript grows to the length of 5- or 6-mer, the transcript would

clash with σR3.2. This structural data was reaffirmed by biochemical assays in which it was shown RNAP lacking σR3.2 do not synthesise abortive transcripts (Murakami et al., 2002b). Consequently, following clash of the nascent transcript with σR3.2, it is thought the nascent transcript either dislodges σR3.2 from its position, or the short RNA is released as an abortive transcript. Along with DNA scrunching stress, this mechanism is thought integral to RNAP promoter escape and the transition to elongation. To escape the promoter, sequence specific interactions within the open promoter complex must be broken. This principle is supported by the observation that, somewhat paradoxically, promoter escape is negatively correlated with the strength of the promoter (Ko and Heyduk, 2014, Hsu et al., 2006, Vo et al., 2003), supporting the proposed idea energy created within the initiation complex determines the breaking of sequence specific contacts. By clashing with and displacing σR3.2 the growing transcript displaces σR4 from the initiation complex, enabling RNAP to escape the promoter, dissociate from σ^{70}, and allow transition to processive elongation.

1.3 RNA polymerase targeting antibacterials

The essential nature of RNAP and its divergence from eukaryotic counterparts make it an excellent target for antimicrobial therapy. A highly complex molecular architecture means a large number of antibacterials, both natural and synthetic, bind RNAP and inhibit various stages of the transcription cycle. Additionally, conserved sequence homology amongst prokaryotic RNAPs often facilitates broad spectrum activity amongst RNAP targeting antibacterials. Furthermore, RNAP inhibitors are not just valuable antibiotics, they are often useful aids with which to dissect RNAP function. Known RNAP inhibitors can generally be grouped into a several main categories depending on the mechanism by which they inhibit RNAP; (i) inhibitors of nascent RNA extension (ii) inhibitors of holoenzyme assembly (iii) nucleoside analogues (iv) inhibitors of active centre mobile elements (v) inhibitors of NTP uptake (vi) inhibitors of open complex formation.

1.3.1 Inhibitors of nascent RNA extension

Ansamycins (Rifamycins)

Ansamycins are a family of natural product antibiotics produced by several strains of Actinobacteria (Kim et al., 2006, Wang et al., 2012). They possess a distinctive structure, comprised of an aromatic nucleus, commonly a naphthoquinone moiety, bridged at its non-adjacent ends by an aliphatic chain (Figure 1.9). A subclass of the ansamycins, the rifamycins, were first isolated in 1959 from the fermentation broth of *Amycolotopsis mediterranei,* and found to possess potent antibacterial activity on account of strong inhibition of prokaryotic RNAP (Sensi et al., 1959). The principal rifamycin, rifamycin B, was only moderately active, but could be further converted into two more potent species, rifamycin O and rifamycin S (Sensi et al., 1960) (Figure 1.9). Reduction of rifamycin S to rifamycin SV (Rif SV) yielded the first drug compound of the class to be introduced to the clinic (Figure 1.9). Rif SV was used intravenously and topically in the treatment of *staphylococcal* infections, hepatic infections, and *tuberculosis* infections. However, despite good antibacterial activity, Rif SV exhibited poor oral bioavailability and marked liver sequestration (Floss and Yu, 2005). Subsequently, Rif SV became the basis of extensive structure activity relationship (SAR) experimentation with the aim of resolving pharmacokinetic problems.

Early SAR studies indicated hydroxyl groups at C1 and C8 on the naphthoquinone core, and at C21 and C23 on the *ansa-* bridge were essential for rifamycin function. Modification of any of these functional groups rendered the compound inactive (Bacchi et al., 1998). Furthermore, these essential oxygen functionalities must remain in a specific spatial orientation for activity (Bacchi et al., 2008). Consequently, most alterations to the *ansa*-bridge, which generally effect the conformation of the aliphatic bridge, abolish or significantly reduce activity. The naphthalene moiety was identified as the most amenable to synthetic modification, particularly at positions C3 and C4. Early efforts focused primarily around modifications at C3 (Floss and Yu, 2005). Addition of a 4-methyl-1-piperazinyl moiety at C3 yielded by far the most important and widely used clinical rifamycin, rifampicin (Rif) (Figure 1.9). The compound was introduced to

the clinic in 1968 and remains the front-line treatment against *Mycobacterium tuberculosis* infections, the causative agent of tuberculosis (Aristoff et al., 2010).
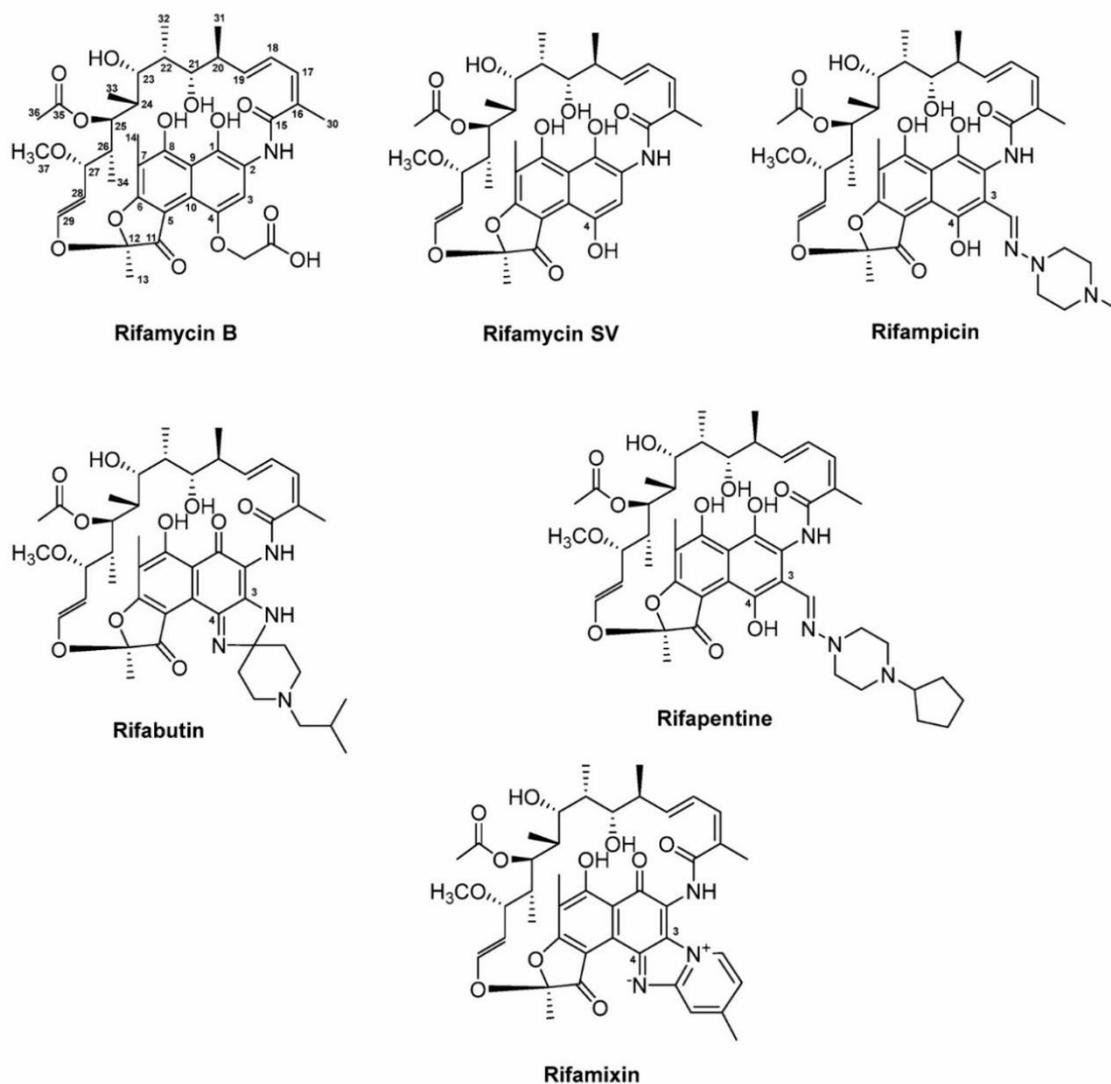
Figure 1.9 Chemical structures of natural and semi-synthetic rifamycins. Adapted from (Robertsen and Musiol-Kroll, 2019). Structures of the naturally derived rifamycins; rifamycin B and rifamycin SV, and the semi-synthetic derivatives rifampicin, rifabutin, rifapentin, and rifamixin.

Fascinatingly, Rif is 1000 times more potent at mycobacterial RNAPs, when compared to *E. coli* RNAP. This difference is not thought to be due to differences in binding site, as this region of RNAP is highly conserved in RNAPs (Zenkin et al., 2005a). Indeed, the mechanistic details of this phenomenon remain elusive. Several further analogues of Rif, rifabutin, rifapentine, and rifamixin were

synthesised by further modifications focused around the C3/C4 positions, and are approved for the treatment of a broad range of infections (Aristoff et al., 2010) (Figure 1.9). Rifamycins are generally broad spectrum antimicrobials and exhibit their highest activity against gram positive bacteria, with MICs commonly below 0.1µg/ml (Aristoff et al., 2010).

The antibacterial activity of rifamycin is due to potent inhibition of DNA-dependent RNA synthesis as a consequence of binding to prokaryotic RNAP (Campbell et al., 2001, McClure and Cech, 1978). Co-crystal structures of several rifamycins bound to RNAP show the inhibitors bind at a site within the main RNAP cleft, around 12 Å from the active centre (the 'Rif pocket') (Campbell et al., 2001, Artsimovitch et al., 2005). They bind in the pathway of nascent RNA and sterically block its propagation beyond 2-3nt in length (Figure 1.10). When the extending RNA collides with the bound rifamycin, it is released as a short abortive product. If RNA extends beyond 3-mer, rifamycins cannot bind to RNAP. Consequently, rifamycins are unable to inhibit the elongation complex (McClure and Cech, 1978, Campbell et al., 2001).

This 'steric occlusion' mechanism, broadly speaking, is mutual to all rifamycins. However, slight mechanistic differences exist depending on the substituent present at C3/C4. For example, rifabutin inhibits the formation of the first phosphodiester bond (when transcription is initiated with a 5' non-phosphorylated dinucleotide primer), whereas Rif and rifapentine inhibit synthesis of the second phosphodiester bond (Artsimovitch et al., 2005). It was proposed an additional 'allosteric mechanism' may act in addition to the 'steric occlusion' mechanism, in which rifamycins allosterically modulate RNAPs affinity for catalytic Mg^{2+}. However, this model was later invalidated (Feklistov et al., 2008). Indeed, most biochemical and structural data support 'steric occlusion' as the principle mechanism of the rifamycins.
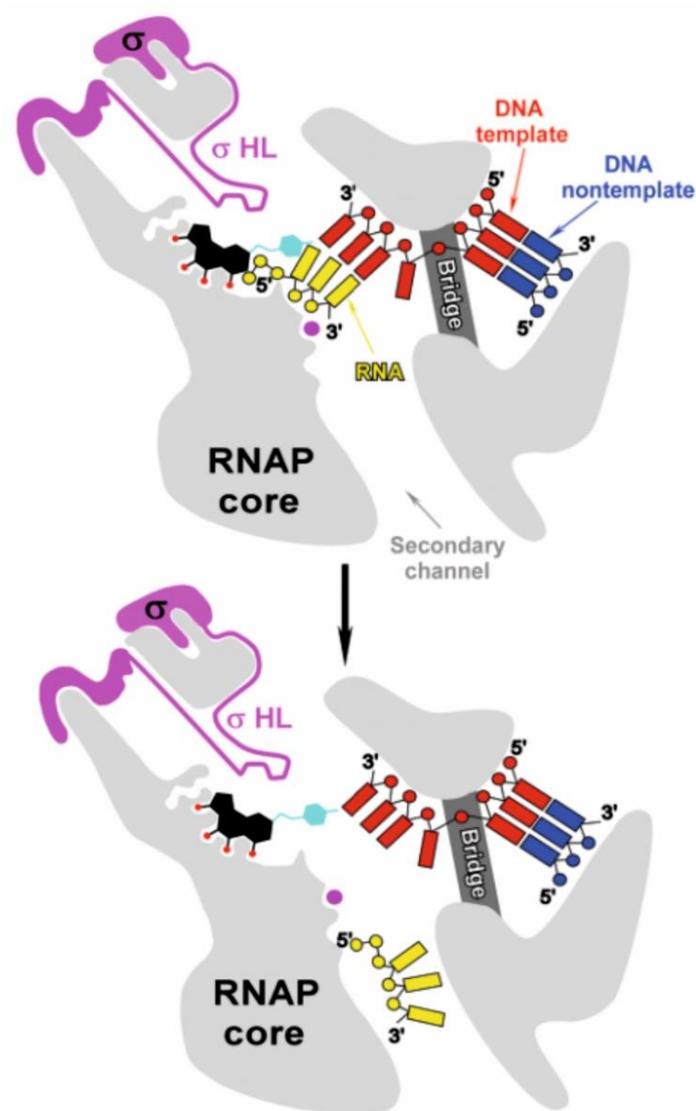
Figure 1.10 Mechanism of transcription inhibition by rifampicin. Adapted from (Artsimovitch and Vassylyev, 2006). An initially transcribing complex (upper panel) composed of core enzyme (light gray, with bridge helix (bridge) shown in dark gray), σ factor (magenta, with σR3.2 (σHL) shown extending into the primary channel toward the RNA:DNA hybrid), promoter DNA (template strand in red and non-template strand in dark blue), and nascent RNA (yellow) base paired with the template strand. Catalytic Mg^{2+} is depicted in magenta. Binding of Rif (black/blue) causes a steric clash with the 5' triphosphate of nascent RNA causing it its dissociation through the RNAP secondary (lower panel)

Resistance to Rif (Rif^R) is most commonly caused by amino acid substitutions of the Rif-pocket that alter residues involved in binding contacts, or alter the conformation of the binding pocket, thus preventing binding of the antibiotic (Campbell et al., 2001, Molodtsov et al., 2017a). As shown in Figure 1.11, the majority of Rif^R mutations found on the β subunit are grouped into a region

termed the Rif-resistance determining region (RRDR), which is further subdivided into four distinct clusters; cluster I (amino acid residues 507-534, *E. coli* numbering), cluster II (amino acid residues 563-572), cluster III (amino acid residues 684-690), and the N-terminal cluster (amino acid residues 143-146) (Goldstein, 2014, Tupin et al., 2010b) . Almost ~100 unique amino acid substitutions have been identified in clinical isolates of Rif[R] *Mycobacterium tuberculosis (*Sandgren *et al., 2009).* Interestingly, however, only three of these mutations account for the vast majority of Rif[R] clinical isolates*,* with βS531L, βH526 and βD516V mutations accounting for ~41%, 36% and 9%, respectively (Campbell et al., 2001) (Figure 1.11). The rapid selection rate for mutations (10^{-7}-10^{-8}) of the Rif pocket observed with rifamycins remains a major drawback in their use and generally limits there use to extensive combination therapies (Goldstein, 2014).
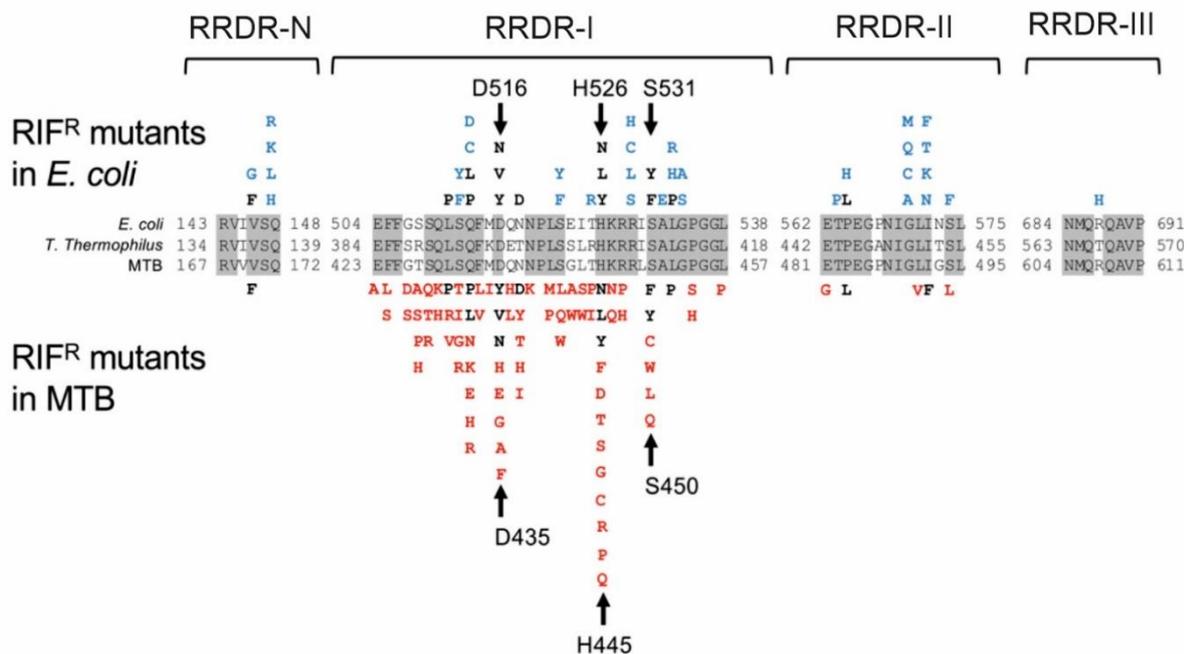
Figure 1.11 Sequence alignment of rifampicin resistance determining regions (RRDRs) with common mutations observed in *E. coli* and *M. tuberculosis (*MTB) conferring resistance to rifampicin. Adapted from (Molodtsov et al., 2017a). Individual RRDR regions are indicated above the sequence alignments. Positions that share the same amino acid are depicted with a gray background. Mutations that confer resistance in *E. coli* or MTB are indicated, with major mutation sites indicated. Mutations exclusive to *E. coli* RNAP are depicted in blue, mutations exclusive to MTB RNAP are depicted in red, and mutations found in both RNAPs are depicted in black.

In addition to alteration of RNAP, resistance to rifamycins is known to occur through a number of other mechanisms, including target duplication, antibiotic modification, and decreased cell permeability. Genome sequencing of the Rif resistant pathogen *Nocardia farcinica* IFM 10152 identified the presence of two genes encoding the RNAP β subunit, *rpoB1* and *rpoB2 (*Ishikawa et al., 2004). Further analysis of the two genes revealed *rpoB2* possesses amino acid substitutions in the RRDR that likely result in resistance to Rif. Indeed, generation of *rpoB2* knockout strains confirmed the gene confers Rif resistance. This mechanism of gene duplication is apparently common among species of *Nocardia* and has also been described in a related *Actinomadura* strain (Vigliotta et al., 2005), suggesting Rif producing *Actinomadura*, such as *A. mediterranei,* may utilise this mechanism to resist the antibiotic. However, at present, this theory is unconfirmed.

Inactivation of rifamycins by covalent modification is a strategy utilised by a number of different bacteria, and is thought to prevent binding of the antibiotic to RNAP. Bacteria can inactivate Rif in a number of different ways, including hydroxylation, glycosylation, phosphorylation, and ADP-ribosylation. In addition to duplication of the β subunit, *N. farcinica* encodes a Rif monooxygenase which catalyses 2' N-hydroxylation of Rif, resulting in a compound with greatly reduced antibacterial activity (Abdelwahab et al., 2016). Additionally, 2' N-hydroxylation of Rif is thought to prime the compound for degradation. A related monoxygenase has also been identified in the pathogen *Rhodococcus equi (Andersen et al., 1997).* Certain Bacillus strains, Nocardia species and related actinomycetes are also able inactivate Rif by O-23 glycosylation and O-21 phosphorylation (Dabbs et al., 1995, Tanaka et al., 1996). Furthermore, a recent phosphotransferase isolated from *Listeria monocytogenes* phosphorylates Rif at the C21 position (Stogios et al., 2016). The opportunistic pathogen *Mycobacterium smegmatis* inactivates Rif by ADP-ribosylation at the C23 position, a reaction catalysed by rifampicin ADP-ribosyltransferase enzymes (Arr) (Morisaki et al., 2000). Rifamycin SV, and newer semi-synthetic Rif derivatives rifapentine, rifamixin and rifabutin, are also substrates for Arr (Baysarowich et al., 2008, Shin et al., 2018)*.* Homologues are widely distributed in environmental bacteria and also present in certain gram negative pathogens such as *P. aeruginosa, Klebsiella pneuominiae*

Actinetobacter baumannii, and certain *E. coli* strains (Shin et al., 2018, Houang et al., 2003, Arlet et al., 2001, Naas et al., 2001, Tribuddharat and Fennewald, 1999). Interestingly, Rif analogues with bulky carbamate substituents at the C25 position are resistant to inactivation by *M. smegmatis* Arr (Combrink et al., 2007). Indeed, recently solved structures of Rif complexed with *M. smegmatis* Arr suggest the bulky substituents would prevent Rif binding to the enzyme, consequently evading ADP-ribosylation (Baysarowich et al., 2008).

A further strategy of resistance to rifamycins (although not exclusive to rifamycins) is through altered membrane permeability. For antibiotics with cytoplasmic targets the compounds must enter the cell and accumulate at sufficiently high concentrations for activity. Bacteria can prevent this in two ways; (i) by an intrinsic or acquired decrease in membrane barrier permeability, and (ii) by overexpression of membrane associated efflux pumps. Mycobacteria possess an inherently impermeable cell envelope as a consequence of their waxy, mycolic acid rich cell wall. Early investigations of a Rif resistant Mycobacterium, *Mycobacteria intracellulare*, showed that the strain was intrinsically resistant to Rif, yet its RNAP was sensitive to the antibiotic (Hui et al., 1977). Further experimentation showed resistance was due to inability of the compound to cross the cell envelope. Several efflux pumps are implicated in Mycobacterial resistance to Rif and several other antibiotics. The putative efflux pump Tap is upregulated by certain Rif resistant Mycobacteria when grown in the presence of Rif (Szumowski et al., 2013), suggesting expression of efflux pumps can be induced by Mycobacteria in response to treatment with Rif.

Sorangicin

Sorangicin (Sor) is a macrolide polyether antibiotic isolated from the fermentation broth of *Myxobacterium cellulosum*. It is a complex of two active structural variants, A and B (Irschik et al., 1987). Sor exhibits broad spectrum antibiotic activity but is predominantly effective against Gram positives, particularly Mycobacteria (Irschik et al., 1987*)*. Sor was shown to effectively inhibit both *E. coli* and *T. aquaticus* RNAP at IC_{50}'s below 1 µM (Campbell et al., 2005). The crystal structure of Sor complexed with *T. aquaticus* RNAP showed the inhibitor binds RNAP within the Rif binding pocket (Campbell et al., 2005) (Figure 1.12 A,

B). Indeed, the mechanism of Sor is essentially the same as Rif; the compound sterically blocks formation of the second phosphodiester bond, causing the nascent transcript to be released as a short abortive RNA (Campbell et al., 2005). Due to the mutual binding site with Rif, there is considerable, but not comprehensive, cross resistance (Xu et al., 2005, Campbell et al., 2005). Molecular dynamics simulations indicate Sor possesses greater conformational flexibility within the Rif binding pocket, and consequently is less susceptible to alterations in binding pocket structure. This flexibility is thought to allow Sor to retain activity at some Rif^R RNAPs. For example Sor retains activity at the prevalent Rif^R mutation βS531L (*E. coli* numbering) (Campbell et al., 2005). However, it is worthy to note, selection of spontaneous mutations conferring resistance to Sor are selected for at a rate almost identical to Rif (Rommele et al., 1990).
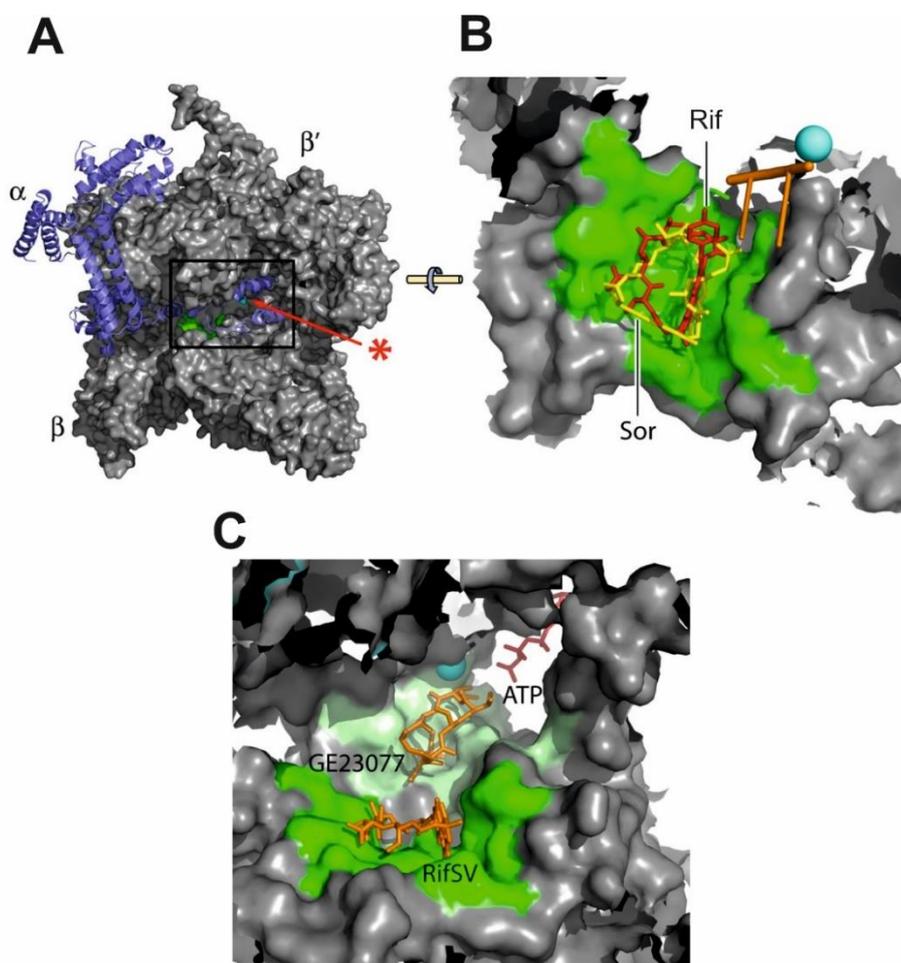
Figure 1.12 Binding site of rifampicin, sorangicin, and GE23077 on RNAP. Adapted from (Ma et al., 2016) (A) *T. thermophillus* holoenzyme with core subunits (α, β, and β') depicted in gray, and σ initiation factor in blue. The catalytic Mg^{2+} is depicted in cyan, and its position indicated by the

red asterisk and arrow. The Rif/Sor binding site is depicted in green. The boxed region is depicted in B and C. (B) The overlapping binding site of Sor and Rif. Initiating RNA dinucleotide is depicted in orange. Rifampicin is depicted in red and sorangicin is depicted in yellow. Other structures in view are depicted as in A. (C) The binding site of GE23077 (orange). The compound binds in the i and i+1 sites (pale green). ATP entering the active centre through the secondary channel is depicted in red. Other structures in view are depicted as in A. Proximity to the Rif binding site (green) is illustrated by bound Rif SV (orange). A hybrid molecule of Rif SV covalently GE23077 remains active against WT and Rif^R RNAPs.

GE23077

GE23077 (GE) is a macrocyclic heptapeptide antibiotic isolated from the culture of *Actinomadura* species in 2004 (Ciciliato et al., 2004). The compound is a specific inhibitor of both WT and Rif^R bacterial RNAP (Sarubbi et al., 2004, Zhang et al., 2014a). However, it exhibits limited antibacterial activity due to limited membrane permeability (Sarubbi et al., 2004, Zhang et al., 2014a, Ciciliato et al., 2004). Activity is improved against *ΔtolC* strains, and when GE is combined with a membrane perturbing agent. GE inhibits transcription initiation by preventing the synthesis of 2nt nascent transcripts. The crystal structure of GE complexed with *Thermus thermophilus* RNAP holoenzyme showed the compound binds the i and i+1 sites within the RNAP active centre, adjacent to the catalytic Mg^{2+} (Figure 1.12, C). Numerous binding contacts are formed with residues of the β and β' subunits, including the three aspartate residues responsible for coordination of catalytic Mg^{2+} (Zhang et al., 2014a). By binding within the i and i+1 sites, GE acts to sterically hinder the binding of initiating NTP substrates required to start transcription. Interestingly, despite establishing extensive binding contacts, only four residues of the β subunit (βE565, βG566, βM681 and βN684) have been identified that confer viable resistance to GE, suggesting the i and i+1 sites offer a promising target for antibiotics. Furthermore, the proximity of the GE binding site to the Rif pocket has enabled the synthesis of a bipartite molecule combining Rif SV and GE that remains active against WT and Rif^R RNAPs (Zhang et al., 2014a) (Figure 1.12, C). By producing such hybrid antibiotics it is hoped the rate of selection of resistance mutations can be reduced, and pharmacokinetic issues addressed.

1.3.2 Disrupting holoenzyme assembly

SB-2 series

Transcription initiation requires the binding of a σ initiation factor to RNAP core to facilitate formation of the initiation competent holoenzyme (Ruff et al., 2015). A high throughput screen of the ChemBridge^{TM} library, in which binding affinity between *E. coli* RNAP core and the housekeeping σ^{70} was measured, identified a series of synthetic compounds termed the SB-2 series (André et al., 2006, Andre et al., 2004). Two initial hit compounds, SB11 and SB15, were identified for their specific inhibition of core : σ^{70} association. Subsequently, both compounds were found to inhibit *in vitro* transcription performed by *E. coli* RNAP holoenzyme. A series of furanyl rhodanine derivatives of the SB series were shown to possess antibiotic activity against a number of Gram positive bacteria, specifically from the *Bacillus, Streptococcus* and *Staphylococcus* genera (Villain-Guillot et al., 2007b, André et al., 2006).

Fascinatingly, SB series compounds are capable of inhibiting σ-independent transcription at poly (dA-dT) DNA template, suggesting the binding site of SB compounds is situated on RNAP core (André et al., 2006). Furthermore, this observation complicates the apparent mechanism of action by which SB compounds function. It is likely the compound inhibits transcription by RNAP holoenzyme by hindering interactions between σ and RNAP core. The compounds inhibit transcription following the formation the holoenzyme suggesting their hindrance of σ: core interactions is mediated allosterically (André et al., 2006). However, it is also possible the compound targets a structural element of catalytic core required for function. However, as of yet, no structure of bacterial RNAP complexed with SB compounds has been solved. Mutations conferring resistance to SB series compounds also remain elusive. Consequently, the SB series binding site and exact mechanism of action remains unclarified.

1.3.3 Nucleoside analogues

Pseudoridmycin

Nucleoside analogues are compounds structurally related to NTP substrates. They can often compete with NTPs for their respective binding sites and consequently inhibit nucleic acid binding enzymes, including nucleic acid polymerases (Périgaud et al., 1992). Pseudoridmycin (PUM) was the first nucleoside analogue identified capable of specific inhibition of bacterial RNAP (Maffioli et al., 2017). It is produced by several bacteria from the *Streptomyces* genus (Rosenqvist et al., 2019, Maffioli et al., 2017). Initially, the compound was identified from a large screen of microbial extracts for its ability to inhibit *E. coli* RNAP, but not the structurally unrelated RNAP of the SP6 bacteriophage. The compound displayed antibacterial activity against a wide range of bacteria, including an array of drug-resistant pathogens (Maffioli et al., 2017).

Inhibition of bacterial transcription by PUM *in vitro* indicated the compound inhibits RNAP by competing with UTP for the i+1 site. Incorporation of UTP, but not ATP, CTP, or GTP, was inhibited by PUM in both single and multiple nucleotide addition experiments. Furthermore, PUM lost activity at DNA templates that do not direct incorporation of UTP. Spontaneous mutations conferring resistance to PUM in *E. coli* mapped to a region within the RNAP active centre, overlapping the i+1 site (Maffioli et al., 2017). Indeed, the crystal structure of *T. thermophilus* RNAP complexed with PUM confirmed PUM targets the i+1 site (Figure 1.13, A,B). The compound forms Watson-Crick base pair interactions with residues of the template strand and forms polar interactions with several residues of the i+1 site in a manner analogous to NTP substrates (Maffioli et al., 2017) (Figure 1.13, B, C). The formation of Watson Crick interactions between the PUM base moiety and template strand are only possible at template positions directing incorporation of UTP, offering an explanation for PUMs targeting of UTP incorporation.
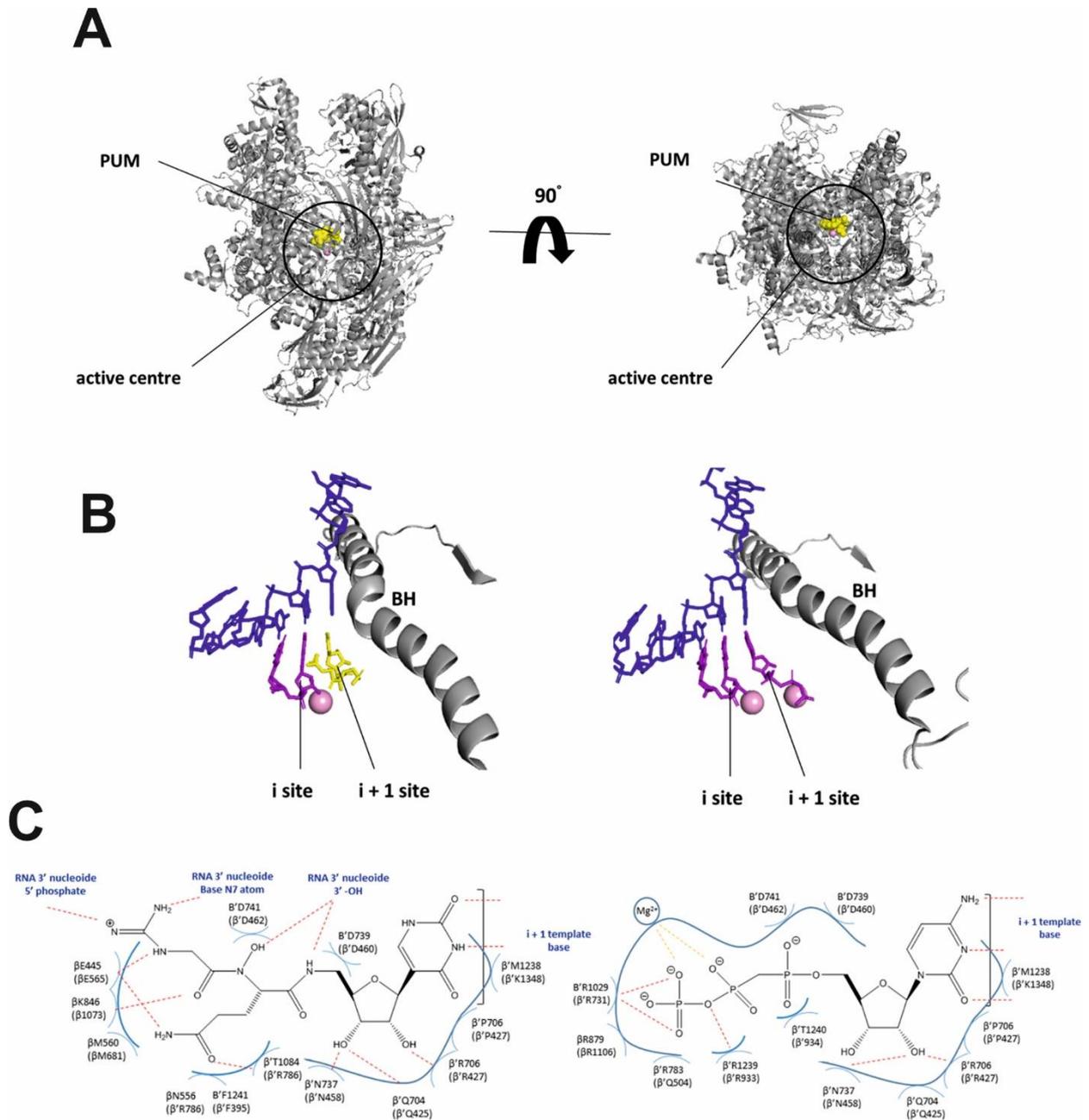
Figure 1.13 Inhibition of RNAP by Pseudoridmycin (PUM). Adapted from (Mosaei and Harbottle, 2019). (A) Orthogonal views of PUM binding position within the active centre of *T. thermophilus* RNAP. RNAP is depicted in gray, PUM is depicted in yellow, and the active centre Mg^{2+} is depicted as magenta sphere. (B) The active centre of T. thermophillus RNAP with PUM (yellow) (left) and CMPcPP (purple) (right); a non-hydrolysable nucleoside, bound within the active centre i+1 site. The RNAP bridge helix (BH) is depicted in gray, RNAP active centre Mg^{2+} depicted as magenta spheres, and template DNA is depicted in blue. (C) Vital interactions of PUM (left) and CMPcPP (right) within the i+1 site of *T. thermophilus* RNAP (*E. coli* numbering in brackets). Note the comparable interactions of PUM and CMPcPP binding.

1.3.4 Inhibitors of mobile elements of the active site

Salinamides

The salinamides are a group of structurally related compounds belonging to a rare class of depsipeptides isolated from several marine *Streptomyces* species (Trischman et al., 1994, Moore et al., 1999). Salinamide A (Sal), B and F are effective antibiotics as a consequence of potent inhibition of bacterial RNAP (Hassan et al., 2015, Degen et al., 2014). Biochemical experiments showed Sal inhibits both transcription initiation and elongation, as well as the reverse reaction of phosphodiester bond formation - pyrophosphorolysis. Sal does not compete with binding of NTP substrates, nor does the compound inhibit the formation of the promoter complex (Degen et al., 2014). By identifying spontaneous and induced Sal resistance mutations, *Degan et al* identified a putative binding site adjacent to the RNAP active centre, overlapping the N-terminal hinge of the BH (BH-H_N), plus two other structural components implicated in the conformational changes of the BH during nucleotide addition, the 'F-loop' and 'link region'. Indeed, Cocrystalisation of Sal and *E. coli* RNAP showed the inhibitor binds in a region between the secondary channel and BH, making direct interactions with the BH-H_N, link region and F-loop (Figure 1.14 A, B). The compound interacts with the BH-H_N in an unbent (straight) conformation (Degen et al., 2014). A bent conformation of BH-H_N has been identified in several molecular dynamics simulations, and is thought to be an important intermediate during catalysis of both phosphodiester bond formation and pyrophosphorolysis (Kireeva et al., 2012, Weinzierl, 2010). Thus, it is thought Sal inhibits RNAP by stabilising BH-H_N in an unbent conformation and consequently prevents conformational cycling of the BH-H_N between the bent and straight conformations important for catalysis. Additionally, the crystal structure of RNAP complexed with Sal possessed a disordered TL (Degen et al., 2014). Structural modelling of Sal binding within the crystal structure of a *T. thermophilus* RNAP elongation complex bound to NTP substrate suggests Sal may sterically hinder correct folding of the TL. Thus, Sal inhibition of TL conformational changes may contribute to the mechanism by which Sal inhibits transcription. Although, biochemical data presented by Degen et al suggested Sal inhibition was not dependent on TL, a recent smFRET experiment in which TL conformations were monitored during nucleotide addition

indicated Sal inhibits TL folding in solution, suggesting TL is indeed a target for Sal (Mazumder et al., 2019, Degen et al., 2014).

Streptolydigin

Streptolydigin (Stl) was first isolated from *Streptomyces lydicus* in 1955 (Lewis et al., 1955, Crum et al., 1955, Deboer et al., 1955). It possessed a distinctive structure containing an acylated tetramic acid, a 'streptolol' side chain, and a monosaccharide moiety. Stl exhibited potent antibacterial activity against a broad spectrum of bacteria by inhibiting bacterial RNAP. The compound was able to inhibit initiation, elongation, and pyrophosphorolysis activities (Temiakov et al., 2005b). Additionally, early biochemical studies showed Stl inhibits both translocation and nucleotide binding (McClure, 1980). Interestingly, spontaneous Stl resistant mutations had been reported in two distinct clusters adjacent the RNAP active centre, β543–546 and β'792–793 (Tuske et al., 2005, Yang and Price, 1995, Severinov et al., 1995). The two clusters occupied distinct non adjacent regions, approximately 15Å apart, located close to the active centre. Saturation mutagenesis experiments delineated 72 individual mutations from 26 distinct residues, defining a determinant that overlapped three distinct structural features; the BH, TL, and a region of the β subunit termed the 'Stl pocket' (comprised of residues β543–545 and β570–571) (Tuske et al., 2005, Vassylyev et al., 2007b, Temiakov et al., 2005a). Interestingly, some of the determinant slightly overlapped the proposed binding determinant of the transcription inhibiting lasso peptide Microcin J25 (see 1.3.5) (Braffman et al., 2019).

Indeed, several crystal structures of *T. thermophilus* RNAP complexed with Stl show the inhibitor binds a site adjacent to, but not overlapping, the RNAP active site (Temiakov et al., 2005a, Tuske et al., 2005, Vassylyev et al., 2007b) (Figure 1.14, A, B). The monosaccharide moiety of Stl occupies a space proximal to the TL and central region of BH, making polar interactions with downstream DNA and hydrophobic interactions with the TL. The streptolol moiety interacts with the 'Stl pocket' and the N terminal end of the BH, whilst the tetramic acid moiety interacts directly with the TL, resulting in its displacement. Interestingly, deletion of the TL increases Stl binding affinity at *T. thermophilus* RNAP (Temiakov et al., 2005a). It is proposed, as a consequence of its binding mode, Stl acts to trap the BH in a

straight conformation, whilst concurrently trapping the TL in an unfolded conformation. In doing so, Stl stabilises the RNAP active site in an inactive substrate bound conformation in which accurate substrate loading is disfavoured. Consequently, the inhibitor blocks the dynamic conformational changes essential for effectual catalysis (Temiakov et al., 2005a).

CBR Series

The original CBR series compound, CBR703, was identified in a high throughput screen of synthetic compounds in which activity against *E. coli* RNAP was assessed, and activity against *E.coli ΔtolC* determined (Zhu et al., 2014, Artsimovitch et al., 2003). The CBR series possess a distinctive structure comprised of two linked aromatic rings. *In vitro* analysis showed CBRs inhibited all catalytic activities of RNAP, but had minimal effect on translocation of RNAP (Artsimovitch et al., 2003). Interestingly, CBR703 was active against RNAPs from a number of Gram-positive and Gram-negative species, but failed to inhibit *Mycobacterium tuberculosis RNAP* (Feng et al., 2015). Recent crystal structure of RNAP complexed with CBR703 show the compound binds at a hydrophobic two-pocket site adjacent to the β' fork loop and N-terminal region of the bridge helix, in addition to the β subunit link domain, DII loop, and F-loop 2 (Figure 1.14). The binding mode is supported by identification of several spontaneous and induced mutations conferring resistance to CBR703. Interestingly, the selection rate of CBR703 resistance mutations in certain *E. coli* strains is as low as 1×10^{-12}, indicating the CBR binding site may be a desirable target for future therapeutic antibiotics (Feng et al., 2015).

Inhibition of RNAP by CBR703 is thought to be caused by allosteric inhibition of TL folding, as consequence of interaction with the β' fork loop, and through inhibition of conformation cycling of the BH-H_N (Bae et al., 2015b) (Figure 1.14, C). It has also been proposed CBR compounds weaken BH-TL interactions, consequently destabilising TL folding required for catalysis (Malinen et al., 2014). Interestingly, however, inhibition of nucleotide addition by CBR703 is only partially dependent upon the TL, whilst inhibition of pyrophosphorolysis occurs entirely independent of the TL. Only inhibition of intrinsic hydrolysis by CBR703 is dependent on TL. Consequently, it is proposed CBR703 targets a previously

unidentified conformational change of the active centre during catalysis of catalytic processes involving NTPs (i.e. nucleotide addition and pyrophosphorolysis). Such conformational changes may be regulated by BH-H$_N$ (Bae et al., 2015b).
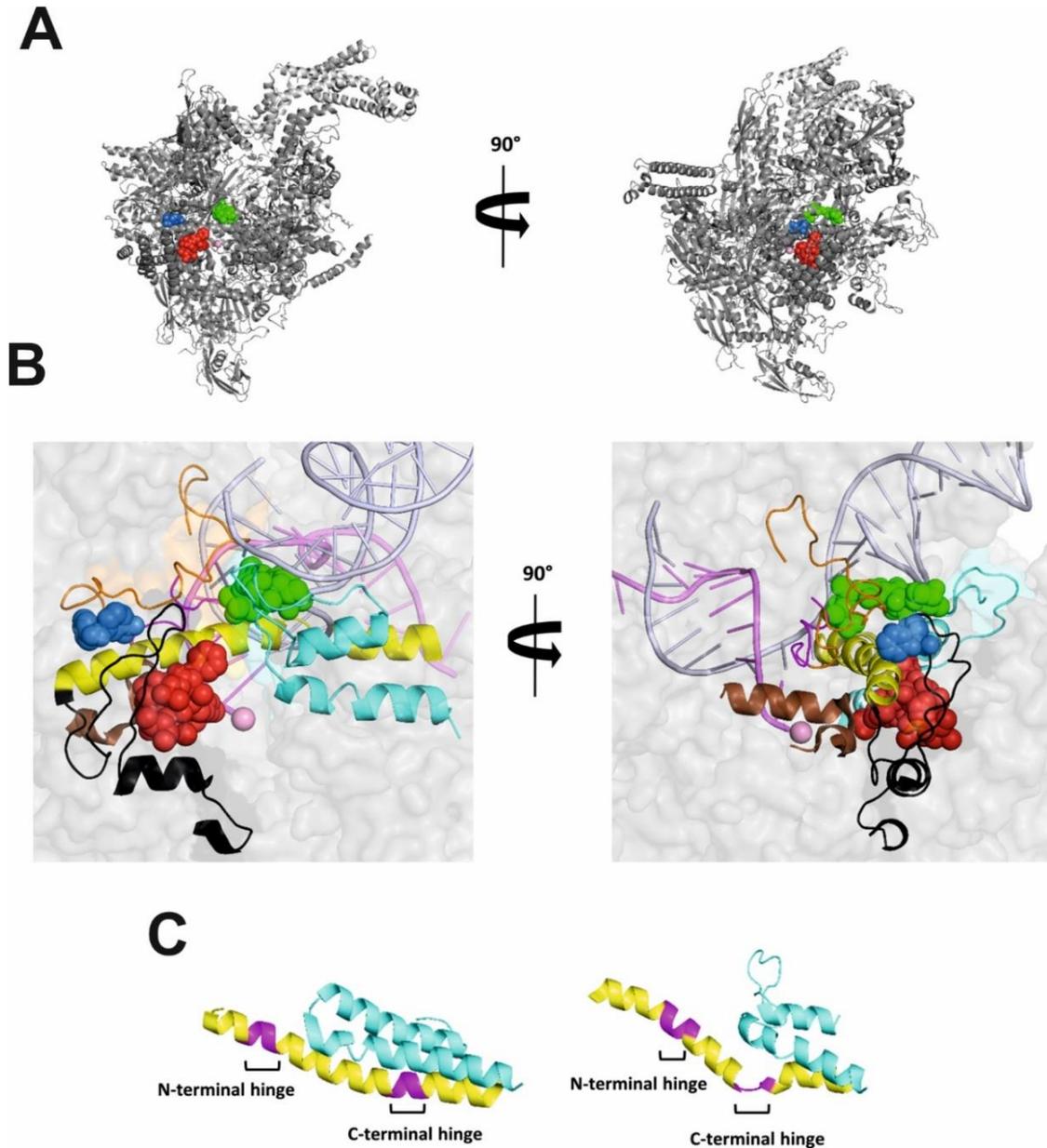
Figure 1.14. Inhibitors of mobile elements of the active centre. Adapted from (Mosaei and Harbottle, 2019) (A) Orthoganal views of binding positions of Sal (Red), Stl (green), and CBR703, (blue) represented in sphere models, mapped onto the structure of *T. thermophilus* RNAP holoenzyme (gray ribbon model). The active centre Mg^{2+} is depicted as a pale pink sphere (B) Close up view of active site mobile elements of *T. thermophillus* RNAP shown in A. RNAP is

shown in a grey semi-transparent surface model and different structural elements of the active centre displayed in ribbon models. The bridge helix (BH) is depicted in yellow, the trigger loop (TL) is depicted in cyan, the 'link' domain is depicted in brown, the DII loop is depicted in magenta, the F-loop is depicted in black, and the F-loop 2 is depicted in orange. Template and non-template DNA is depicted in pale blue and RNA is depicted violet. (C) Structures of different conformations of the BH and TL; closed TL and straight BH associated with closed active site (Left panel); and open TL and bent BH associated with an open active site (right panel). The N-terminal and C-terminal hinges are highlighted in magenta

D-AAP1

A novel class of synthetic compounds was recently identified from the high throughput screen of a synthetic compound library. D-AAP1, the progenitor molecule of the class, was found to possess potent activity against *M. tuberculosis* RNAP *in vitro*, but exhibited poor activity against other bacterial RNAPs and human RNAPs (Lin et al., 2017b). The compound exhibits potent, selective activity against several Mycobacteria including *M. tuberculosis, M. avium,* and *M. smegmatis*, but poor activity against other bacterial and mammalian cells. The crystal structure of *M. tuberculosis* RNAP in complex with D-AAP1 and isolation of resistance determinants demonstrated the inhibitor binds a pocket adjacent to the BH-H$_N$ (Lin et al., 2017b). The binding site overlaps directly with that of CBR series compounds. Consequently, the two compounds are thought to inhibit RNAP through the same mechanism. However, the specificity of D-AAP1 for Mycobacteria is drawn from subtle structural differences seen in Mycobacterial RNAP. The binding site of D-AAP1 is comprised of three individual protein pockets on the surface of *M. tuberculosis* RNAP which interact with the three ringed structure of the compound. On the other hand, the respective binding site of CBR series compounds on *E. coli* RNAP is comprised of two protein pockets on the enzyme surface which interact with the two ringed structure of CBRs (Lin et al., 2017b).

1.3.5 Inhibitors of NTP uptake

Microcin J25

Microcin J25 (MccJ25) is a cyclic 21 residue antibiotic peptide synthesised by *E. coli* strains containing the pTUC plasmid encoding the gene cluster *mcjABCD* (Bayro et al., 2003, Wilson et al., 2003). It was first identified in 1992 from *E. coli* strain AY25 (Salomón and Farías, 1992). MccJ25 possesses a distinctive

threaded lasso structure with 2 distinct elements; a 'lariat ring' formed through a covalent linkage between the N-terminus and glutamic acid at position 8, through which a 'tail' structure, comprised of residues 9-21, is sterically trapped (Bayro et al., 2003, Wilson et al., 2003) (Figure 1.15, A). MccJ25 inhibits both transcription initiation and elongation of *E. coli* RNAP *in vitro* (Adelman et al., 2004). Mutations conferring resistance to MccJ25, generated through saturation mutagenesis, mapped almost entirely to the secondary channel, indicating a putative binding site (Mukhopadhyay et al., 2004, Yuzenkova et al., 2002). It was proposed MccJ25 acts to plug the RNAP secondary channel through a so called 'cork-in-a-bottle' mechanism, consequently stopping uptake of NTP substrates into the active centre. Certainly, MccJ25 increases K_m of NTP binding, supporting the proposed mechanism (Mukhopadhyay et al., 2004). However, many MccJ25 resistance determinants overlap the binding site of the active site inhibitor Streptolydigin on the β subunit, indicating certain elements of their mechanisms of action may be shared (Yuzenkova et al., 2002). Indeed, binding of MccJ25 and streptolydigin occurs competitively, further substantiating this proposal.

The recently published crystal structure of MccJ25 complexed with *E. coli* RNAP confirmed the peptide binds deep within the secondary channel (Braffman et al., 2019) (Figure 1.15, B). The tail structure is oriented towards the active centre, with the lariat ring located roughly 6.5 Å from the catalytic Mg^{2+}. Modelling of MccJ25 binding within *T. thermophilus de novo* initiation complex suggests the peptide would afford both an electronic and steric clash with the triphosphate moiety of the 3'-NTP substrate. Furthermore, MccJ25 reduces the solvent accessible gap of the secondary channel from 11 Å to below 5 Å, decreasing the accessibility of NTP substrates to the active centre (Braffman et al., 2019). These observations offer an explanation for the increased K_m for NTP binding seen with MccJ25. In addition to obstruction of NTP substrate binding, MccJ25 forms binding interactions with the BH and unfolded TL. When bound to RNAP the peptide introduces a severe steric clash to correct TL folding (Braffman et al., 2019) (Figure 1.15, C, D). As a result, TL folding is highly disfavoured in the presence of MccJ25. Correct folding of the TL is essential for closure of the active site during nucleotide addition. Consequently, it appears MccJ25 inhibits RNAP through a tripartite mode of action; i) preventing cyclic conformational changes of

the TL vital for efficient catalysis, ii) obstructing access of NTP substrates through the secondary channel, and iii) impeding binding of NTPs within the active site.
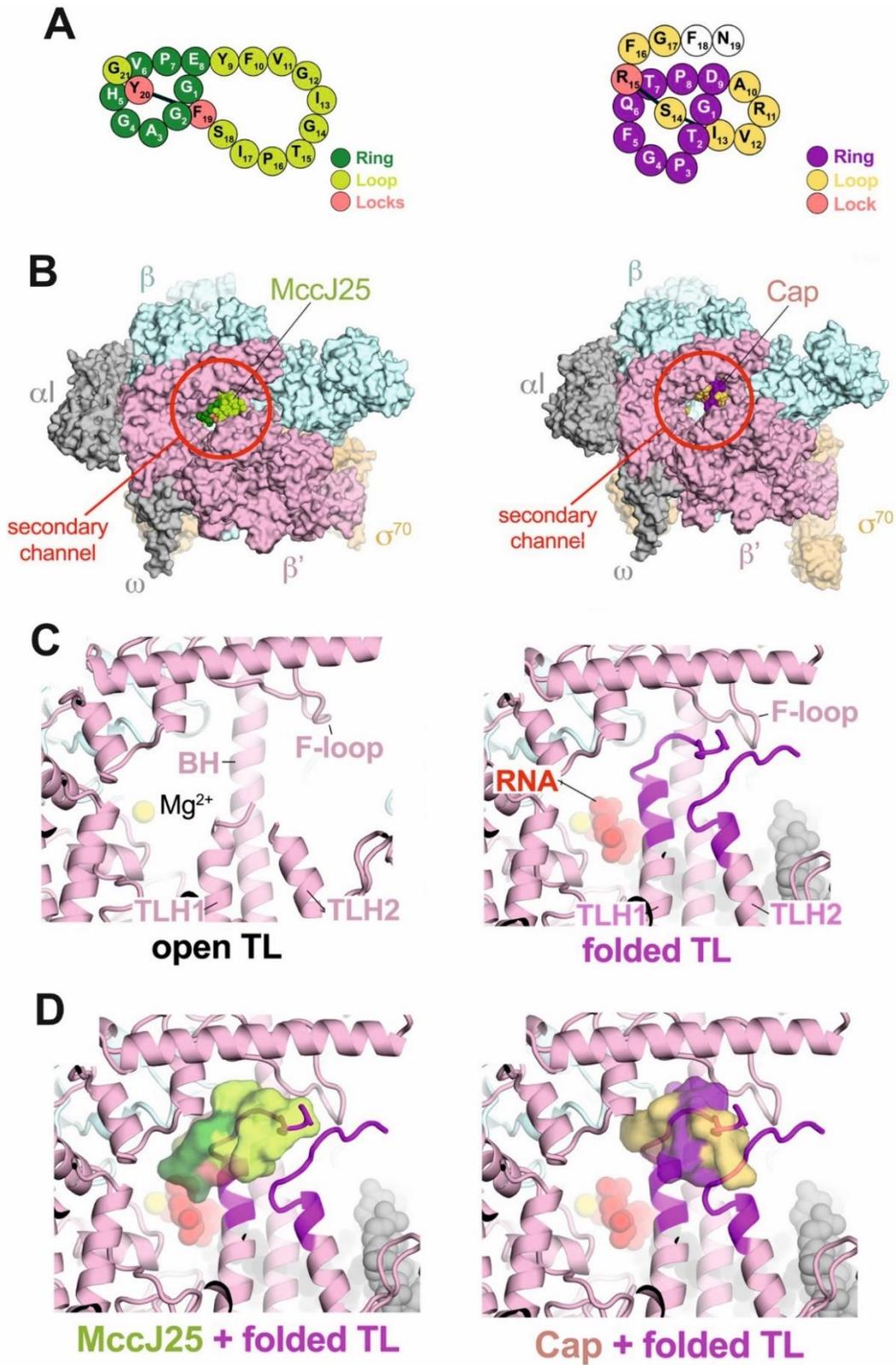
Figure 1.15 Inhibition of RNAP by lasso peptide microcin J25 (MccJ25) and capistruin (Cap). Adapted from (Braffman et al., 2019). (A) Schematic of MccJ25 (left) and Cap (right). (B) Overall structure of *E. coli* RNAP holoenzyme complexed with MccJ25 (left panel) (molecular model coloured as in A) and Cap (molecular model coloured as in A) RNAP is shown as a surface model with subunits coloured as in their respective labels. (C) View into the secondary channel with varying trigger loop conformations of *E. coli* RNAP. RNAP is depicted as a ribbon representation (β, cyan; β', light pink) The Active centre Mg^{2+} is depicted as a yellow sphere, the two alpha helices or the TL are depicted TLH1 and TLH2. The F-loop and BH structures are indicated.   . (Left panel) Structure of *E. coli* RNAP with an open (unfolded) TL. (Right panel) Structure of the *E. coli* RNAP transcription initiation complex with a closed (folded) TL. Nucleic acids are shown as molecular sphere models (DNA, light gray; post translocated RNA transcript, red). (D) Both panels are as right panel of C, with binding site of MccJ25 (left) and Cap (right) superimposed. Note the steric clash with the closed TL induced by both lasso peptides.

Capistruin

Capistruin (Cap) is a 19 amino-acid lasso peptide produced by *Burkholderia thailandesis* E264 (Knappe et al., 2008)*.* It was identified from a genome mining experiment aimed at identifying putative lasso peptide precursors, and homologs of Mcc25 processing enzymes. Cap is structurally comparable to MccJ25; its N-terminus is covalently fused to an aspartate residue a position 9, producing a 9 amino-acid ring, through which the C-terminal tail threads, sterically locked in position by an arginine residue at position 15 (Knappe et al., 2008) (Figure 1.15, A). Cap was shown to inhibit *in vitro* transcription by WT E. coli RNAP, but failed to inhibit an MccJ25 resistant RNAP containing a β'T931I mutation (Kuznedelov et al., 2011). This indicated Cap may bind at the same site on RNAP. Indeed, the crystal structure of *E. coli* RNAP complexed with Cap shows the peptide also binds within the secondary channel (Braffman et al., 2019). However, despite Cap sharing several binding interactions with residues involved in Mcc25 interactions, Cap binding determinants are distinct. The peptide binds RNAP with a similar orientation to MccJ25; ring and tail proximal to the active site, and loop distal (Figure 1.15, B). Yet, Cap binds at a site further from the active centre, within a region of the secondary channel that is wider than that occupied by MccJ25. Consequently, Cap does not obstruct the channel to the same degree as MccJ25, and would seemingly not preclude access of NTP substrates. Furthermore, Cap resides almost 12 Å away from the active site Mg^{2+}. Modelling within *T. thermophilus de novo* initiation complex indicates Cap and NTP substrates can easily be accommodated simultaneously, indicating Cap inhibition

is not caused by preventing access and binding of NTPs to the active site. Indeed, biochemical analysis showed transcription inhibition by Cap is principally not dependent on NTP concentration (Braffman et al., 2019). However, like MccJ25, Cap interacts with both the BH and an unfolded TL. When bound to RNAP the peptide introduces a severe steric clash to correct TL folding in a manner analogous to Mcc25 (Figure 1.15, C, D). Therefore, in the presence of Cap, folding of the TL is highly disfavoured. Hence, despite a similar binding mode to MccJ25, it seems the principal mode of action of Cap is through inhibition of correct TL folding, a process essential for efficient catalysis of nucleotide addition by RNAP.

1.5.6 Inhibitors of promoter open complex formation

Fidaxomicin (Lipiarmycin)

Fidaxomicin (Fdx), also known as lipiarmycin, is a first in class macrocyclic antibiotic first identified from culture of the actinomycete bacteria *Dactylosporangium aurantiacum* (Johnson, 2007). It has recently been approved for clinical treatment of *clostridium difficile*-associated diarrhoea. Initial biochemical experiments indicated Fdx targets the RNAP switch region, inhibiting promoter melting and σ-dependent transcription initiation (Tupin et al., 2010a, Lin et al., 2017a). The compound doesn't inhibit the formation of the promoter complex, but was shown to inhibit the binding of template DNA within the RNAP active centre (Tupin et al., 2010a). Two recent studies utilising cryo-electron microscopy (cryo-EM) elucidated the structure of Fdx complexed with *M. tuberculosis* RNAP (Lin et al., 2017a, Boyaci et al., 2018). The compound interacts with SW2, SW3 and SW4, in addition to several clamp alpha helices. Of the 5 discrete switch elements, SW1 and SW2 are the principal mediators of conformational changes of the clamp. The inhibitor makes 5 essential polar interactions with several residues of β and β' (β'Q94, β'R99, β'248, β'337 and βK1303; *E. coli* numbering). Substitutions at any of these positions confer resistance to Fdx (Lin et al., 2017a).

Structural data indicated the binding of Fdx to RNAP locks the clamp domain in an *open* conformation (Lin et al., 2017a, Boyaci et al., 2018) (Figure 1.16). Consequently, it is thought Fdx prevents the correct spatial orientation of the

clamp required for simultaneous engagement of both the -10 and -35 promoter elements by σR2 and σR4, respectively. This observation is reaffirmed by biochemical observations indicating Fdx-RNAP complexes can bind to upstream promoter elements, but fail to engage the -10 promoter element, and therefore fail to nucleate promote melting (Lin et al., 2017a, Morichaud et al., 2016, Tupin et al., 2010a). Specifically, it is proposed the 'tryptophan wedge' (Trp wedge), responsible for intercalation into the NT strand at the -12 position to nucleate promoter melting, cannot engage DNA when the clamp is in an open conformation (Figure 1.16, B). Additionally, the protein pocket on σR2 responsible for binding and stabilising the flipped out non template base at the -11 position, cannot interact with DNA with the clamp locked open (Lin et al., 2017a). Therefore, Fdx is thought to inhibit σ-dependent transcription initiation by preventing recognition of the -10 promoter element by σR2, and consequently RNAP cannot transition from RPc to RPo.
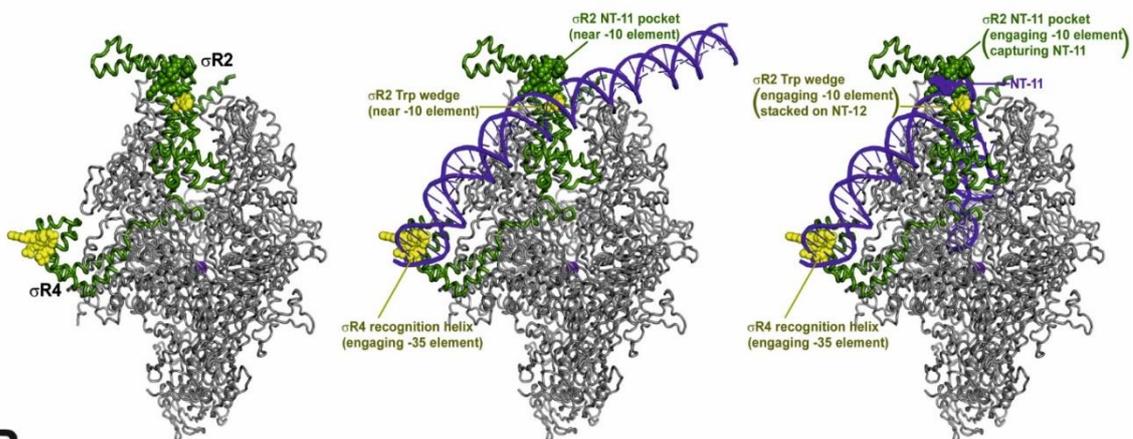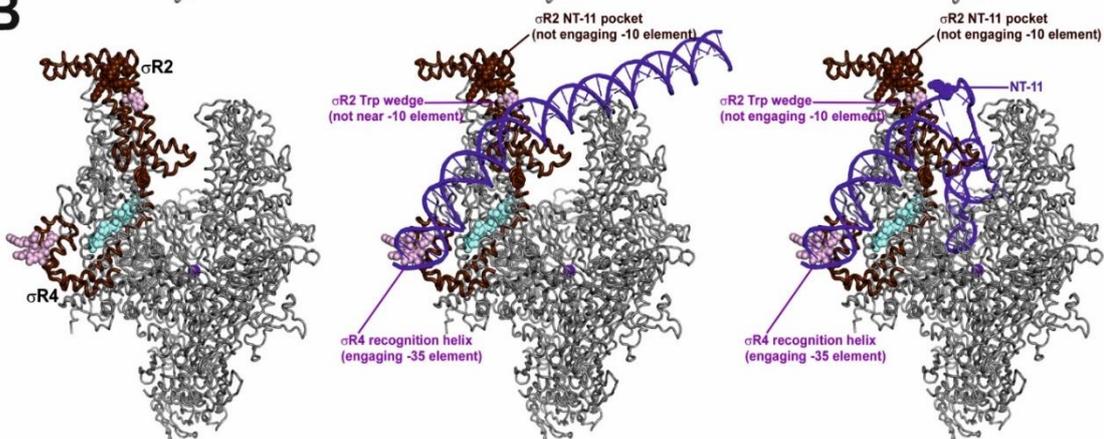
Fig 1.16 Mechanism of action of Fidaxomicin (Fdx). Adapted from (Lin et al., 2017a). (A) Structure of M. tuberculosis RNAP holoenzyme with a closed clamp conformation in the absence of Fdx (left), RPc (centre) and RPo (right. σ is depicted in green ribbon model, σR2 is depicted in green surface model, non-template -11 pocket is depicted as yellow surface model, σR2 'Trp wedge' depicted as a blue surface model, σR4 is depicted as a yellow surface, and DNA is depicted as a blue cartoon model. Note the ability of σR2 and σR4 to simultaneously engage promoter −10 and −35 elements, respectively, in both RPc and RPo. (B) Same as A, except Fdx (cyan) is bound to RNAP with clamp locked in an open conformation. Colours are as in A, except σ is depicted as a brown ribbon model, σR2 NT-11 pocket is depicted as a brown surface model, σR2 Trp wedge and σR4 recognition helix are depicted as a pink surface model. Note in the presence of Fdx, σR2 and σR4 cannot simultaneously engage promoter −10 and −35 elements in RPc and RPo.

Squaramides, myxopyronin, corallopyronin and ripostatin

Myxopyronin (Myx), corallopyronin (Cor) and ripostatin (Rip) are switch region targeting natural products isolated from several species of Myxobacteria (Schäberle et al., 2014). The inhibitors bind within a pocket adjacent to the Fdx binding site and inhibit RNAP through a slightly different mechanism (Boyaci et al., 2019, Mukhopadhyay et al., 2008, Belogurov et al., 2008, Srivastava et al., 2011). Structural analysis of *T. thermophilus* RNAP complexed with Myx indicates the compound interacts predominantly with SW1 and SW2 to lock the RNAP clamp domain in a *closed* conformation (Belogurov et al., 2009, Mukhopadhyay et al., 2008). Introduction of mutations to RNAP elucidated several binding determinants essential for Myx activity, all of which exhibit extensive cross resistance with Cor and Rip, indicating the compounds bind within a mutual pocket (Srivastava et al., 2011, Mukhopadhyay et al., 2008). Indeed, a recent cryo-EM structure of Cor complexed with *M. tuberculosis* promoter complexes showed the compound indeed binds within the same pocket as Myx on RNAP (Boyaci et al., 2019). The structure of RNAP-Cor promoter complexes elucidated a putative intermediate *enroute* to RPo, in which the promoter is partially melted upstream of the transcription start site (Figure 1.17). Indeed, biochemical analysis of Myx and Cor show the compounds do not entirely prevent promoter melting, but instead prevent the propagation of promoter melting reaching the TSS (Srivastava et al., 2011, Mukhopadhyay et al., 2008). The Cor-RNAP structure indicates late promoter melting might occur within the RNAP active centre cleft, and the inhibitors likely act to trap a late intermediate by locking the clamp in a *closed* conformation. The structure

indicates a transient opening of the clamp is essential for propagation of melting to the TSS, owing to restricted access to the active site cleft due to confined space between fork loop-2 and SW2 when the clamp is in a *closed* conformation (Boyaci et al., 2019). Furthermore, binding of both Cor and Myx cause a refolding of SW2 that is thought to induce a large steric clash with the position of template DNA in RPo. This observation is reaffirmed by inhibition of transcription by Myx and Cor at artificially melted promoters (Srivastava et al., 2011). Therefore, it appears Myx, Cor and Rip inhibit transcription by preventing essential opening of the RNAP clamp during promoter melting, and also act to prevent correct positioning of template DNA within the active site.
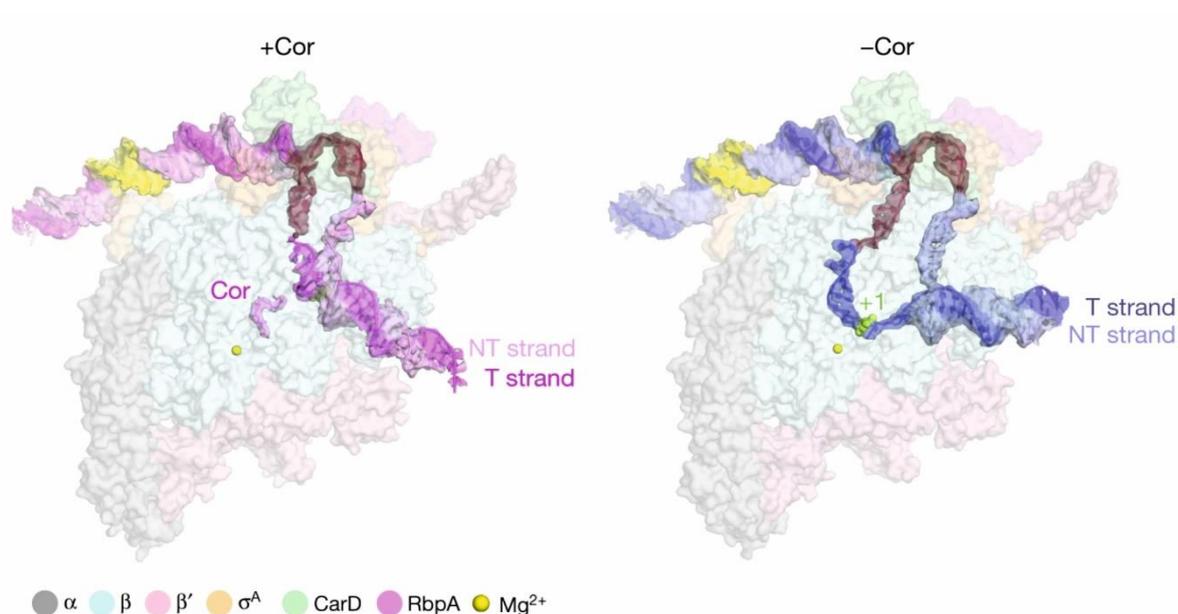
Figure 1.17 Inhibition of RNAP by Corallopyronin (Cor). Adapted from (Boyaci et al., 2019) (Left panel) Structure of *M. tuberculosis* RNAP promoter complex in the presence of Cor. Colouring of structures is as indicated by the figure legend. Cor traps a putative promoter melting intermediate with the promoter partially melted upstream of the TSS. Note template DNA is not loaded within the active site. (Right panel) Structure of *M. tuberculosis* RNAP promoter complex in the absence of Cor. Note the fully melted promoter DNA with template loaded within the active site.

Squaramides (SQ) are a class of synthetic compounds that also target the RNAP switch region (Molodtsov et al., 2015). The compounds are active against efflux negative strains of *E. coli* and *Haemophilus influenzae* (Buurman et al., 2012). Crystal structures of SQ complexed with *E. coli* RNAP show the inhibitors bind within the same pocket as Myx and Cor. SQ acts to displace SW2 in a similar manner to Myx and Cor, likely effecting clamp conformation, and leading to a steric clash with template DNA at positions +3 and +4 (Molodtsov et al., 2015). Therefore, it is likely SQ inhibit RNAP is a similar manner to Myx, Cor and Rip, by sterically occluding binding of melted promoter DNA within the active site, whilst simultaneously locking the clamp in a closed or partly closed conformation.

Chapter 2. Materials & Methods

2.1 Reagents and Antibiotics

All chemicals, antibiotics and reagents were purchased from Sigma unless otherwise stated. All enzymes and their respective buffers were purchased from New England Biolabs. NTPs, chromatography columns, and phosphorimaging screens were purchased from GE Healthcare. Ureidothiophene was purchased from ChemBridge[TM]. All promoter DNA fragments were produced by PCR using Phusion DNA polymerase from their respective primers (IDT) and purified by agarose gel electrophoresis (Qiagen). All radiochemicals were purchased from Hartmann Analytic.

2.2 PCR

PCR was performed in 50 µL reactions containing the following; 200 µM dNTPs, 1 unit of Phusion DNA polymerase, ~ 1 ng of template DNA, 10 µM forward primer and 10 µM reverse primer and 10 µl 5X Phusion High Fidelity Buffer. All primers used can be found in the appendix. The cycling parameters of each individual reaction were as recommended by the manufacturer (New England Biolabs). Annealing temperature was varied depending upon primer T_m and extension time set at 30 seconds per 1000 amplified base pairs.

2.3 Growth Media and Strains

All bacterial strains used in this work are described in the Appendix. Unless otherwise stated, all strains were grown in liquid Luria – Bertani (LB) medium (1 % tryptone, 0.5 % yeast extract, 1 % NaCl) or plated on to LB agar (1% tryptone, 0.5% yeast extract, 1% NaCl, 2% agar) and supplemented with the relevant antibiotic when required. Ampicillin was added to a final concentration of 100 µg/ml. Kanamycin was added to a final concentration of 50 µg/ml.

2.4 Actinomycete Extract Preparation

All actinomycete bacteria, provided by Demuris[TM], were grown on GYM agar (0.4 % glucose, 0.4 % yeast extract, 1 % malt extract and 1 % agar) for roughly 7 days at 30 °C. Growth medium was then extracted with methanol (MeOH) and evaporated to dryness under negative pressure to yield aqueous extract. Extracts

were then cleaned up on a C18 HyperSep™ solid phase extraction (SPE) cartridge (ThermoFisher Scientific) and eluted with MeOH.

2.5 Tandem Liquid Chromatography – Mass Spectrometry

All analytical separations were performed on an Agilent 1260 HPLC by injection of 1-5 μl of sample onto a Raptor ARC-18 LC-2.7 μm - 150×2.1 mm column (Restek) or an Ultra C4 5 μm 150 x 2.1 mm operated at 0.2 μl/min and eluted using a 30 min linear gradient from 5 % to 100 % acetonitrile. Mobile phases were supplemented with 0.1 % formic acid. Mass spectra were recorded in positive-ion mode on a Bruker MicrOTOF II time-of-flight mass spectrometer.

2.6 Disk diffusion assay

Standard disk diffusion assay (Kirby-Bauer) was performed with respective strain as described (Bauer et al., 1966). Briefly, paper disks were loaded with up to 10 μg of compound, air-dried and placed on LB plates with an embedded lawn of respective strain. Reporter strains disk assays were performed with X-gal infused agar (100 μg/ml). Plates were incubated overnight at 37°C and scanned.

2.7 Molecular Cloning

DNA inserts for cloning were generated by PCR from genomic DNA of respective strain. Genomic DNA was purified using the GenElute™ Bacterial Genomic DNA Kit (Sigma) by the manufacturer's protocol. Primers used can be found in the appendix. Amplified insert DNA was purified using the Qiagen Gel Extraction kit, following the protocol provided. Plasmid and insert DNA were restricted using the appropriate restriction enzymes for 1 hour at 37 ^{0}C before being gel purified. 1 unit of Thermosensitive Alkaline Phosphatase (Promega) was added to the digested plasmid and incubated for a further 30 mins. For ligation, 3 times molar excess of insert DNA to plasmid DNA was incubated at rt using T4 DNA ligase (New England Biolabs) for 1-2 h. 5 μl of ligation mix was transformed into 50 μl of DH5α competent *E. coli* cells (New England Biolabs). Cells were transformed according to the supplied protocol. Cells were plated on LB containing the respective selection antibiotic, either 100 μg/ml ampicillin or 50 μg/ml kanamycin. Roughly 5 colonies were picked and grown overnight in LB supplemented with selection antibiotic at aforementioned concentrations. Plasmid DNA was isolated

using the Qiagen Mini Prep Kit. Presence of insert was then assessed by PCR amplification. Individual PCR reactions were loaded onto a 1% agarose gel to resolve the amplified insert. Positive clones were then sequenced to verify insert presence. All sequencing was carried out by Eurofins Genomics using their standard primers. Resultant sequences were visualised using SnapGene software.

2.8 Site-directed mutagenesis

Amino acid substitutions were introduced using the QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies). PCR was performed in 50 μL reactions containing the following; 125 ng forward primer, 125 ng reverse primer, ~ 10 ng of template DNA, 200 μM dNTPs, 5 μl of 10X reaction buffer, 3 μl of QuikSolution, 2.5 units PfuTurbo DNA Polymerase. All primers used can be found in the Appendix. The cycling parameters of each individual reaction were as recommended by the manufacturer's instructions. Annealing temperature was varied depending upon primer T_m.

Following PCR, 10 units of DpnI (Agilent Technologies) was added to each reaction for 1 hour at 37 °C to digest parental DNA. 2 μl of the DpnI treated reaction were transformed into DH5α competent *E. coli* cells (New England Biolabs) according to the manufacturer's guidelines.

2.9 Buffers

All Buffer compositions can be found in respective methods text. All chromatography buffers were filtered through bottle top 0.45 μm PVDF filters (Sarstedt) prior to use.

2.10 Protein expression and purification

2.10.1 Purification of *E. coli* core RNAP

Core *E. coli* RNAP subunits were expressed in T7 express cells (New England Biolabs) transformed with pGEMABC (encoding rpoA, rpoB, and rpoC) and pACYCDuet-1_Ec_rpoZ (encoding rpoZ) (Murakami, 2013). Expression was induced by addition of 0.4 mM final IPTG to exponentially growing cells and incubated on an orbital shaker (150 rpm) at room temperature overnight. Cells

were then harvested by centrifugation and resuspended in grinding buffer (50 mm
Tris-HCl (pH 7.9), 10% glycerol, 200mM NaCl and 1× protease inhibitor mixture).
Cells were then lysed by sonication and debris cleared by centrifugation. RNAP
was precipitated from the lysate by addition of polyethyleneimine solution to a
final concentration of 0.6% and the pellet recovered by centrifugation. The pellet
was first washed with TGED buffer (10 mM Tris-HCl (pH 7.0), 10% glycerol, 0.1
mM EDTA, and 2 mM DTT) + 0.5 M NaCl. RNAP was then eluted from the pellet
by suspension in TGED buffer 1 M NaCl and then precipitated by ammonium
sulphate to a final concentration of 60 % saturation. RNAP was resuspended in
TGED buffer + 50 mM NaCl. Lysate was filtered through 0.45 µM PVDF filter
(Merck) and injected at 1 ml/min onto a 5 ml HiTrap Heparin Affinity column (GE
healthcare) equilibrated with heparin buffer A (TGED buffer + 50 mM NaCl).
RNAP was eluted by linear gradient to 100 % heparin buffer B (TGED buffer +
1M NaCl) at a flow rate of 1 ml/min, and fractions collected. Individual fractions
were assessed for the presence of RNAP by SDS-PAGE electrophoresis. RNAP-
containing eluates were pooled and injected at 1 ml/min onto a 5 ml Resource Q
ion-exchange column (GE Healthcare) equilibrated with ResourceQ buffer A
(TGED buffer + 50 mM NaCl). RNAP was eluted by linear gradient to 100%
ResourceQ buffer B (TGED buffer + 1M NaCl) at a flow rate of 1 ml/min, and
fractions collected. Individual fractions were assessed for the presence of RNAP
by SDS-PAGE electrophoresis. RNAP-containing fractions were pooled and
concentrated by centrifugation by Ultra-15 Centrifugal Filter Units (Amicon) at 4ºC
according the manufacturers guidelines. RNAP was then dialysed into storage
buffer (10 mM Tris-HCl (pH 7.5), 50% glycerol, 100 mM NaCl, 0.1 mM EDTA, 1
mM DTT) at 4°C.

2.10.2 Purification of *S. epidermidis* RNAP

Cellular *Staphylococcus epidermidis* WT RNAP holoenzyme and E105Q σ^A
RNAP holoenzyme were purified from respective strains of WT *Staphylococcus
epidermidis* ATCC 12228 *and Staphylococcus epidermidis* ATCC 12228
harbouring an E105Q mutation in *rpoD* (see Isolation of ureidothiophene resistant
staphylococcus epidermidis). Cells were grown to late exponential phase and
then harvested by centrifugation and resuspended in grinding buffer (50 mm Tris-
HCl (pH 7.9), 10% glycerol, 200mM NaCl and 1× protease inhibitor mixture).

Cells were then lysed by sonication and debris cleared by centrifugation. RNAP holoenzymes were purified by HiTrap Heparin Affinity and Resource Q (GE Healthcare) column chromatography essentially as described in 2.10.1.

2.10.3 Purification of mutant *E. coli* RNAPs

RIF-resistant & KglA-resistant mutations were introduced in pIA581 plasmid (Svetlov and Artsimovitch, 2015) (encoding *E. coli* rpoA, rpoB, and rpoC, with 6xHis-tag on N terminus of β subunit) by site-directed mutagenesis. Mutant plasmids were co-transformed with pACYCDuet-1_Ec_rpoZ (encoding rpoZ) (Murakami, 2013) in T7 Express strain (New England Biolabs). Expression was induced by addition of 0.4 mM final IPTG to exponentially growing cells and incubated on an orbital shaker (150 rpm) at room temperature overnight. Cells were then harvested by centrifugation and resuspended in grinding buffer (50 mm Tris-HCl (pH 7.9), 10% glycerol, 200mM NaCl and 1× protease inhibitor mixture). Cells were then lysed by sonication and debris cleared by centrifugation. Lysate was filtered through 0.45 µM PVDF filter (Merck) and injected onto a 5ml His-trap Ni $^{2+}$-NTA column (GE Healthcare) equilibrated with Ni $^{2+}$ buffer A (50 mm Tris-HCl (pH 7.9), 10% glycerol, 600mM NaCl). Mutant RNAPs was eluted by stepwise increase of Ni $^{2+}$ buffer B (50 mm Tris-HCl (pH 7.9), 10% glycerol, 600mM NaCl, 200mM imidazole) to increase imidazole concentration of eluent (0mM, 25mM, 50mM, 100mM and 200mM). Eluates were assessed for the presence of RNAP by SDS-PAGE electrophoresis.

2.10.4 Purification of *E. coli* σ^{70} subunit

E. coli σ^{70} subunit was expressed in T7 express cells (New England Biolabs) transformed with pET28 expression vector encoding N-terminal 6x His-tagged *E. coli* σ^{70} subunit. Expression was induced by addition of 0.4 mM final IPTG to exponentially growing cells and incubated on an orbital shaker (150 rpm) at room temperature overnight. Cells were then harvested by centrifugation and resuspended in grinding buffer (50 mm Tris-HCl (pH 7.9), 10% glycerol, 200mM NaCl and 1× protease inhibitor mixture). Cells were then lysed by sonication and debris cleared by centrifugation. *E. coli* σ^{70} was then purified by HisTrap HP (GE Healthcare) nickel affinity chromatography essentially as described in 2.10.3. Ureidothiophene resistant mutations were introduced in pET28a plasmid

encoding N-terminus 6x His-tag *E. coli* σ^{70} by site-directed mutagenesis (Agilent technologies) and purified as above.

2.10.5 Purification of *M. smegmatis* and *M. abscessus* Arr

M. smegmatis and *M. abscessus Arr* was expressed in T7 express cells (New England Biolabs) transformed with pET28 expression vector encoding N-terminal 6x His-tagged *M. smegmatis Arr or M. abscessus Arr*. Expression was induced by addition of 0.4 mM final IPTG to exponentially growing cells and incubated on an orbital shaker (150rpm) at room temperature overnight. Cells were then harvested by centrifugation and resuspended in grinding buffer (50 mm Tris-HCl (pH 7.9), 10% glycerol, 200mM NaCl and 1× protease inhibitor mixture). Cells were then lysed by sonication and debris cleared by centrifugation. *E. coli* σ^{70} was then purified by HisTrap HP (GE Healthcare) nickel affinity chromatography essentially as described in 2.10.3.

2.10.6 Purification of *S. aureus*, *M. smegmatis* and *T. aquaticus* RNAP

M. smegmatis and *T. aquaticus* RNAPs were purified and provided by Dr Amber Riaz-Bradley and Dr Christina Julius (Newcastle University) as described (Kuznedelov et al., 2003, Mukherjee and Chatterji, 2008). *S. aureus* RNAP was provided by Dr Caitlin Griffiths (Newcastle University).

2.11 5' radiolabelling of RNA and DNA primers

5'-radiolablled DNA primers were used to synthesise 5'-labelled template DNA by PCR (as described in 2.2). Where applicable, 5' radiolabelled RNA primers were used in the assembly of artificially assembled elongation complexes. For 5' labelling, 25 µL reactions containing the following; 2.5 µl Primer (10 µm final), 1 unit T4 Polynucleotide Kinase, 5 µl γ-[^{32}P]-ATP (10 mCi/ml) and 2.5 µl. 10x PNK A Buffer incubated at 37 °C for 1 hour. The reaction was inactivated by heating to 75ºC for 10 mins and cleaned up on a bio-spin 6 column (Bio-Rad).

2.12 In vitro transcription assays

2.12.1 In vitro transcription on promoter DNA

Transcription from promoter DNA fragments was performed essentially as described (Mosaei et al., 2018). Briefly, reactions were performed in 10 µL of transcription buffer TB (20 mM Tris HCl pH 7.9, 40 mM KCl, 10 mM $MgCl_2$). 1 pmol of *E. coli* RNAP core with 3 pmols of σ^{70} or 1 pmol of *T. aquaticus* RNAP core with 3 pmols of *T. aquaticus* σ^A or 1 pmol of *M. smegmatis* or *S. aureus* or *S. epidermidis* RNAP holoenzymes were incubated in TB with 1 µL of DMSO (or 50% MeOH in the case of bacterial extracts) containing or not containing inhibitor at 37°C (or 60°C in case of *T. aquaticus* RNAP) for 5 mins. Transcription was initiated by the addition of 2 µL mixture of nucleotides and promoter DNA in TB, containing (final concentrations): 10 nM promoter DNA, 25 µM CpA (for T7A1 and GalP1 promoters) or 100 µM ApA (for *lac*UV5 promoter), 0.2 µl α-[^{32}P]UTP (10mCi/ml) (Hartmann Analytic), 10 µM UTP with (run off transcription) or without (abortive transcription) 100 µM ATP, CTP and GTP. Reactions were stopped after 10 min incubation at 37°C (or 60°C in case of *T. aquaticus* RNAP) for run off transcription or 5 minutes for abortive transcription by the addition of equal volume of formamide-containing loading buffer. Products were resolved in denaturing polyacrylamide gels, revealed by PhosphorImaging (GE Healthcare), and analyzed using ImageQuant software (GE Healthcare)

2.12.2 *In vitro* transcription on M13ori hairpins

In vitro transcription on M13ori hairpin template was performed as described in (Zenkin and Severinov, 2004). Briefly, reactions were performed in 20 µL final volumes. 3 pmols of wild-type RNAP core with 15 pmols of σ^{70} and 3 pmols of single-stranded M13ori promoter (IDT) were incubated at 37 °C for 10 min in TB. Transcription was initiated by the addition of 1 mM ATP, CTP and UTP, 100 µM GTP and 0.2 µl α-[^{32}P] GTP (10 mCi/ml) (Hartmann Analytic). Reactions were stopped after 30-min incubation at 37°C by the addition of formamide-containing loading buffer. Products were separated on denaturing polyacrylamide gels, revealed by phosphorimaging (GE Healthcare), and analysed using ImageQuant software (GE Healthcare). All quantifications were made in triplicate.

2.12.3 *In vitro* transcription from artificially assembled elongation complexes

Elongation complexes were assembled as previously described (Yuzenkova et al., 2013). Sequences of assembled elongation complexes used here are illustrated in their corresponding figures. Reactions were carried out in 15 µL final volume. Briefly, RNA was 5'-radiolabelled by T4 polynucleotide kinase and γ-[^{32}P]-ATP prior to complex assembly. The reaction was inactivated by heating to 75°C for 10 mins and cleaned up on a bio-spin 6 column (Bio-Rad). Subsequently, 0.5 pmol of 5'-labelled RNA and 1 pmol template DNA were incubated in TB at 45°C for 5 mins and then cooled slowly to 4°C to anneal the hybrid. The RNA: DNA hybrid was then incubated with 5pmol core RNAP for 5 mins at 37 °C. The complexes were then incubated with 10 pmol non-template DNA bearing a 5'-biotin tag for 5 mins at 37°C .The complexes were then immobilised on 5 µL streptavidin beads and then washed first with high salt (1 M KCl) and then low salt (40 mM KCl) TB. Reactions were then started with one or a combination of 1 µM GTP, CTP, UTP and ATP and incubated at 37 °C for the times indicated in the respective figures. Reactions were stopped by the addition of formamide-containing loading buffer. Products were separated on denaturing polyacrylamide gels, revealed by phosphorimaging (GE Healthcare), and analysed using ImageQuant software (GE Healthcare). All quantifications were made in triplicate. Rate constants were derived from kinetic data curves fitted to a single exponent equation using non-linear regression in SigmaPlot software.

2.13 KMnO$_4$ and DNAse I footprinting

Reactions were performed in 20 µL final volume of TB (20 mM Tris HCl pH 7.9, 40 mM KCl, 10 mM MgCl$_2$). 5 pmol RNAP core and 10 pmols of σ^{70} were incubated in TB. For DNAse I footprinting, 1/10 reaction volume of Urd solution in DMSO was added and incubated at 37°C for 5 mins. An identical volume of DMSO was added to control samples. For KMnO$_4$ footprinting, 1/10 reaction volume of Urd solution in 75% ethanol was added and incubated at 37°C for 5 mins. An identical volume of 75% ethanol was added to control samples. Reactions were supplemented with 0.25 pmol promoter DNA labelled at the 5' end of the non-template strand and incubated for a further 2 minutes at 37°C. Samples were then treated with 0.25 units DNAse I (Roche) or 5mM KMnO$_4$ and incubated at 37 °C for 30 seconds. For DNAse I footprinting, reactions were stopped with equal volume formamide-containing loading buffer. For KMnO$_4$

footprinting, the reactions were stopped with an equal volume of 2-mercaptoethanol. The KMnO$_4$ treated samples were then subject to phenol-chloroform extraction, ethanol precipitated and dried before resuspension in formamide-containing loading buffer. Products were resolved on polyacrylamide gels, revealed by phosphorimaging (GE Healthcare), and analysed using ImageQuant software (GE Healthcare)

2.14 EMSA

Reactions were performed in 20 µL final volume of EMSA Buffer (20 mM Tris HCl pH 7.9, 40 mM KCl, 10 mM MgCl2, 5% glycerol) 2 pmol of RNAP core and 6 pmols of σ^{70} was incubated in EMSA buffer. Next, 1/10 reaction volume of Urd solution in DMSO was added and incubated at 37°C for 5 mins. An identical volume of DMSO was added to control samples. Reactions were supplemented with 0.2 pmol promoter DNA labelled at the 5' end of the non-template strand and incubated for a further 5 minutes at 37 °C. Samples were then treated with 2 µL H$_2$O or H$_2$O with heparin (50µg/ml final) and incubated for a further 2 minutes at 37°C. Samples were then supplemented with 2 µL dye solution (0.25% bromophenol blue and xylene cyanol) and loaded onto 4.5% non-denaturing polyacrylamide gel, revealed by phosphorimaging (GE Healthcare), and analysed using ImageQuant software (GE Healthcare).

2.15 Isolation of ureidothiophene resistant *staphylococcus epidermidis*

Firstly, MIC of *S. epidermidis* ATCC12228 was assessed by serial dilution on a 24 well agar plate (2-fold dilutions from 100µg/ml). Individual wells contained 1 ml solid LB agar supplemented with 2.5% pluronic F68 (Thermofisher). Prior to agar setting, DMSO with or without ureidothiophene was added to the individual well to a final concertation of 5%. *S. epidermidis* ATCC12228 was streaked onto LB agar and grown at 37°C overnight. A single colony was picked and grown in liquid LB to 1 x10^6 CFU/ml. 10 µL of 10^6 CFU/ml *S. epidermidis* inoculant were dotted onto each well and the plate incubated at 37°C overnight. MIC was deduced as the concentration in which no visible cell growth was observed (3.125 µg/ml). Secondly, *S. epidermidis* ATCC12228 was streaked onto standard LB agar and grown at 37 °C overnight. A single colony was picked and grown in liquid LB until

~ 1×10^9 CFU/ml. 100 µL of 10^9 CFU/ml *S. epidermidis* ATCC12228 was streaked onto an LB agar plate containing 4x MIC ureidothiophene (12.5 µg/ml). Ureidothiophene resistant mutants were identified by their growth on this media, and confirmed by re-streaking on the same media. A single resistant strain was identified and sent for full illumina genome sequencing (MicrobesNG). Genome was assembled and SNPs identified by CLC Genomics Workbench software (Qiagen)

2.16 Isolation of Kanglemycin A

Kanglemycin A was purified by Dr Hamed Mosai-Sejzi (Newcastle University) as described (Mosaei et al., 2018)

2.17 X-ray Crystallography

X-ray crystallography data was gathered by Dr Vadim Molodstov (Pennsylvania State) and Professor Katsu Murakami (Pennsylvania State) as described (Mosaei et al., 2018)

2.18 Determination of minimum inhibitory concentrations (MICs) for *M. tuberculosis*

MIC determination data were gathered by Dr Joanna Bacon (Public Health England) as described (Mosaei et al., 2018).

2.19 Rifampicin ADP-ribosyl transferase disk assay

Reactions were performed in 10 µL final volume of Arr Buffer (20 mM Tris-HCl pH 7.9, 40 mM KCl, 0.5 mM $MgCl_2$). *M. smegmatis* or *M. abscessus* Arr at concentration indicated (0µM, 20 µM or 200 µM) was mixed with antibiotic (1 mg/ml final) in 8 µL Arr buffer at 37°C for 5 minutes. 2 µL NAD+ in water was added (10mM final) and incubated for 1h at 37°C. Reaction was quenched with an equal volume of MeOH and spotted onto paper disks and a disk assay performed as described in 2.6 with an embedded lawn of *S. aureus* RM4220.

2.20 *In vitro* Rifampicin ADP-ribosyl transferase activity assay

Reactions were performed in 100 µL final volume of Arr Buffer (20 mM Tris-HCl pH 7.9, 40 mM KCl, 0.5 mM $MgCl_2$). *M. smegmatis* or *M. abscessus* Arr at 10 µM

was mixed with RIF or KglA at 100 µM in 80 µL Arr buffer at 37ºC for 10 minutes. 20 µL NAD+ in water was added (250 µM final) and incubated for 1 h at 37 ºC. Reaction was quenched with 500 µL of Methanol. Methanol was then evaporated under negative pressure and the reaction analysed by LC-MS as described in 2.5.

2.21 Purification of ADP-ribosyl Rifampicin

Reaction was performed in 2000 µL volume of Arr buffer (20 mM Tris-HCl pH 7.9, 40 mM KCl, 0.5 mM $MgCl_2$) containing the following; 25 µM *M. smegmatis Arr*, 5 mg Rifampicin and 20 mM NAD+. Reaction was incubated at 37 ºC for 24 hours and cleaned up on a 25 ml HyperSep™ C8 SPE cartridge. ADP-ribosyl RIF was eluted with 30% MeOH and dried under negative pressure by a HT-6 series 3 evaporator (Genevac) to yield 4.8 mg of ADP-ribosyl RIF. Sample homogeneity was confirmed by tandem LC-MS as described in 2.5.

2.22 Microscale thermophoresis

Binding affinity experiments were carried out on a Monolith NT.115 Series instrument (Nano Temper Technologies GMBH). *M. smegmatis* and *M. abscessus* Arr were labelled with Monolith Protein Labelling Kit RED-NHS 2nd Generation Amine (Nano Temper Technologies GMBH) according to the manufacturers guidelines. Roughly 5 µl of sample in MST buffer (20 mM HEPES (pH 7.9), 40 mM KCl, 10 mM $MgCl_2$) were loaded into Monolith NT.115 premium capillaries and thermophoresis measured for 30 s. Analysis was performed with Monolith software. K_d was quantified by analysing the change in normalized fluorescence (Fnorm = fluorescence after thermophoresis/initial fluorescence) as a function of inhibitor concentration. Curves for K_d data were fitted to a four-parameter logistic equation using non-linear regression in SigmaPlot software.

Chapter 3. Aims

Bacterial RNAP is an excellent target for antibiotics. However, very few clinical antibiotics target RNAP. The growing prevalence of antibiotic resistance amongst pathogenic bacteria demands the identification of novel antibacterial compounds, acting through novel molecular mechanisms. Here, we conduct several distinct projects in which we investigate previously uncharacterised molecular mechanisms underlying inhibition and resistance of transcription targeting antibiotics.

(i)   Most clinical antibiotics are derived from the natural products of actinomycete bacteria. Our industrial collaborators Demuris^{TM} previously compiled a library of actinomycete bacteria that activate an RNAP reporter strain and therefore may produce novel inhibitors of bacterial transcription. Consequently, we aimed to identify and characterise novel inhibitors of bacterial transcription produced by strains from this particular strain library.

(ii)   The synthetic antibiotic ureidothiophene (Urd) was identified within a commercial screen of synthetic compounds in which inhibition of *S. aureus* RNAP was analysed. However, how the compound targets RNAP is unknown. Here, we aimed to characterise the molecular mechanism of action by which Urd inhibits bacterial RNAP.

(iii)   A prior screening program conducted by Demuris^{TM}, and a subsequent collaboration with the lab of Professor Nikolay Zenkin identified the rifamycin type natural product kanglemycin A (KglA) as an inhibitor of rifampicin resistant RNAPs. Here, we aimed elucidate the molecular mechanism of action by which KglA inhibits RNAP

(iv)   Finally, we aimed to characterise ADP-ribosylation of Rif and KglA by *Mycobacterium smegmatis* and *Mycobacterium abscessus* Rifampicin ADP-ribosyltransferase (Arr) enzymes.

By investigating these unique mechanistic processes we aim to further our understanding of how transcription targeting antibiotics function at RNAP, and

understand the exact mechanisms utilised by pathogenic bacteria to facilitate resistance to transcription targeting antibiotics.

Chapter 4. Streptomyces strain DEM40380 produces Antibiotic A39079S-1, a rifamycin type inhibitor of bacterial RNAP

4.1 Introduction

Actinomycete bacteria are prolific producers of bioactive secondary metabolites, many of which are efficacious therapeutics. Roughly two thirds of antibiotics used clinically, including β-lactams, tetracyclines, aminoglycosides, macrolides, rifamycins, and macrocyclic antibiotics are all derived from compounds produced by Actinomycetes. However, very few novel classes of antibiotics have been discovered in the last 40 years (Wohlleben et al., 2016). Research efforts in natural product drug discovery have declined since the 1970s, whilst 'modern methods' of drug discovery, such as rational design, have had limited success in identifying effective antibiotics (Jackson et al., 2018). The emergence of resistance to antibiotics is a growing public health concern, and demands the identification of novel antibacterials with which to treat drug resistant infections (Davies and Davies, 2010). Consequently, research efforts are revisiting actinomycete bacteria with the aim of identifying novel classes of antimicrobials.

RNAP is a validated target for antibiotic therapy, exemplified by the successful use of rifampicin as a front line treatment against *Mycobacterium tuberculosis* infections. However, rapid selection of Rif resistant mutations in the Rif binding pocket remains a major issue, frequently leading to Rif resistant strains of *M. tuberculosis* (Goldstein, 2014). Indeed, there is a pressing need for novel transcription targeting compounds, acting through novel mechanisms, with which to treat resistant pathogens.

Our industrial collaborators Demuris™ possess a unique collection of over 10,000 diverse, highly dereplicated isolates of Actinomycete bacteria, sourced from a variety of terrestrial and marine environments. In an attempt to identify producers of transcription targeting antibiotics, Demuris performed a preliminary screen of their collection in which strains were screened against a *Bacillus subtilus* reporter strain (yvgS) that has *lacZ* fused to the promoter of the bacterial helicase *HelD* (Hutter et al., 2004a). This promoter is upregulated in response to sub-inhibitory concentrations of the transcription inhibitor Rif. Consequently, in a disk assay, it is presumed strain extracts containing specific transcription inhibitors produce a

blue halo at the frontier of the zone of inhibition upon X-gal infused agar plates. By using this reporter strain, Demuris compiled a shortlisted library of 17 actinomycete strains that activated reporter activity of yvgS, and consequently may produce novel inhibitors of RNAP (Figure 4.1).

Here, we analysed this shortlisted library for novel inhibitors of transcription through a two-pronged approach. Individual strain extracts were subjected to *in vitro* transcription assays to identify the presence of selective inhibitors of bacterial RNAP. In parallel, extracts were subjected to tandem liquid chromatography-mass spectrometry (LC-MS), alongside a comprehensive literature search to identify previously characterised RNAP inhibitors. From this approach we aimed to identify producers of novel selective inhibitors of bacterial transcription.

Strain	Genus
DEM10800	*Streptomyces*
DEM10803	*Streptomyces*
DEM1085	*Streptomyces*
DEM1086	*Streptomyces*
DEM10813	*Streptomyces*
DEM10817	*Streptomyces*
DEM10824	*Streptomyces*
DEM10825	*Streptomyces*
DEM10826	*Streptomyces*
DEM10828	*Streptomyces*
DEM10831	*Streptomyces*
DEM10846	*Actinomadura*
DEM33039	*Micromonospora*
DEM40376	*Streptomyces*
DEM40380	*Streptomyces*
DEM40316	*Streptomyces*
DEM40347	*Streptomyces*

Figure 4.1 Table of actinomycete bacterial strains and their respective genus identified by Demuris™ for a screening program aimed at identifying novel inhibitors of bacterial RNAP. Strains were compiled as a result of activation of the *B. subtilus* yvgS reporter strain. The yvgS reporter has *lacZ* fused to the promoter of the bacterial helicase *HelD*; a promoter upregulated in response to sub-inhibitory concentrations of Rif (Hutter et al., 2004b). Consequently, in a disk assay, it is presumed strain extracts containing specific RNAP inhibitors produce a blue halo at the frontier of the zone of inhibition upon X-gal infused agar plates.

4.2 Results

4.2.1 Identification of producers of transcription targeting compounds by *in vitro* transcription analysis and tandem liquid chromatography-mass spectrometry

Firstly, our industrial collaborators at Demuris™ performed a preliminary screen of their extensive actinomycete collection in which strains were screened against the yvgS reporter strain. From this preliminary screen, Demuris™ technicians identified 17 actinomycete strains that activated reporter activity of yvgS, and consequently may produce novel inhibitors of RNAP (Figure 4.1). These strains were subsequently provided to us to pursue the further characterisation described herein.

To identify producers of secondary metabolites targeting transcription, we cultivated individual isolates from the shortlisted library on GYM agar plates for approximately 7 days. Growth mediums were extracted with methanol and evaporated to dryness, from which aqueous extracts were prepared. These crude extracts were subsequently 'cleaned-up' on C18 solid phase extraction (SPE) cartridges and then eluted with methanol in preparation for *in vitro* analysis. Production of active compound(s) by individual strains was confirmed by a disk assay of the respective extract against the *yvgS* reporter strain prior to further *in vitro* analysis.

Following extract preparation, we assessed the effect of extracts on *in vitro* transcription by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. Extracts were added to *in vitro* transcription reactions *before* template DNA. Expectedly, all extracts were able to inhibit WT *E. coli* RNAP, confirming the presence of transcription inhibiting compounds in the respective extracts (Figure 4.2, A). Next, we investigated if this inhibition could be replicated at Rif^R RNAPs. We analysed the effect of extracts in a second *in vitro* transcription assay in which transcription was performed by a mutant RNAP bearing the rifampicin resistant mutation βH526Y. This particular mutation is frequently seen in clinical isolates of rifampicin resistant *Mycobacterium tuberculosis*, and confers high level resistance to Rif (Goldstein, 2014, Mosaei et al., 2018). In this instance, with the exceptions of DEM40376, DEM40380, and DEM40347, all

extracts lost inhibitory activity, suggesting prior inhibition at WT RNAP is likely due to compounds targeting the Rif binding pocket (Figure 4.2, A). Moreover, all extracts that lost inhibitory activity at the mutant RNAP induced a marked increase in transcription levels relative to the control. It is possible these extracts are contaminated with NTP substrates and/or transcription activating compounds that can positively influence transcription rates when rifamycin type inhibitors are redundant at the mutant RNAP.

A
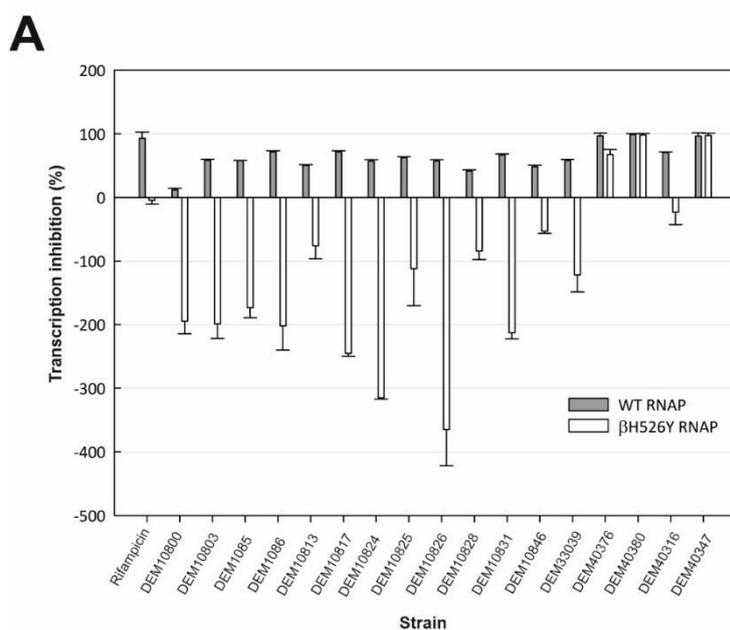
B

Strain	Observed m/z (Da)	Antibiotic adduct
DEM10800	698.2913	[Rifamycin SV+H]
DEM10803	696.2756	[Rifamycin S+H]
DEM1085	756.2569	[Rifamycin B+H]
DEM1086	ND	-
DEM10813	698.2836	[Rifamycin SV+H]
DEM10817	ND	-
DEM10824	698.2887	[Rifamycin SV+H]
DEM10825	698.2814	[Rifamycin SV+H]
DEM10826	ND	-
DEM10828	698.2789	[Rifamycin SV+H]
DEM10831	698.2768	[Rifamycin SV+H]
DEM10846	ND	-
DEM33039	698.2763	[Rifamycin SV+H]
DEM40376	696.2751	[Rifamycin S+H]
DEM40380	ND	-
DEM40316	698.2665	[Rifamycin SV+H]
DEM40347	698.2818	[Rifamycin SV+H]

Figure 4.2 Inhibition of WT and Rif^R *E. coli* RNAP by strain extracts. (A) Quantification of inhibition of *in vitro* transcription by respective strain extracts. Transcription performed by WT *E. coli* RNAP and Rif^R βH526Y RNAP on a linear DNA template containing the T7A1 promoter. Reactions were initiated with dinucleotide primer CpA. Quantification is derived from average of [^{32}P]-labelled run-off and terminated transcription products. Values were normalised to quantity of [^{32}P] RNA synthesised in the absence of inhibitor. Error bars are ± SD from at least 3 independent experiments. (B) Table of exact atomic masses corresponding to known natural product RNAP inhibitors. Masses were identified by tandem liquid chromatography-positive ion mass spectrometry (LC-MS) of respective strain extracts. Subsequently, exact masses [-H] were searched ± 0.05 Da within the dictionary of natural products (http://dnp.chemnetbase.com), and results analysed for respective natural product transcription inhibitors.

To assess the presence of characterised RNAP inhibitors, extracts were subjected to tandem reverse phase liquid chromatography - mass spectrometry (LC-MS). Distinct peaks from total ion chromatograms were identified and corresponding molecular masses searched in the dictionary of natural products, a comprehensive archive of chemical data on natural products. Mass spectrometric analysis confirmed the vast majority of strains produced well characterised rifamycin type inhibitors. Excluding DEM1086, DEM10817, DEM10826, DEM10846, and DEM40380, all strain extracts contained masses corresponding to either rifamycin B, rifamycin S, or rifamycin SV. Of the strains in which no masses corresponding to RNAP inhibitors were identified, DEM1086, DEM10817, DEM10826, and DEM10846, all lost inhibitory activity at rifampicin resistant RNAP. This loss of activity at mutant RNAP suggests rifamycin type inhibitors may be present in these extracts that are hitherto unidentified, or are not yet registered in the dictionary of natural products. Nonetheless, identification and characterisation of novel rifamycins that fail to inhibit rifampicin resistant RNAPs falls outside the scope of this project. Consequently, all strain extracts which contained known rifamycin type inhibitors, and/or failed to inhibit transcription by rifampicin resistant RNAP, were dropped from further analysis. Thus, the strain extract of DEM40380 was selected for further investigation.

4.2.2 DEM40380 extract inhibits both transcription initiation and elongation

Specific inhibitors of RNAP often target a distinct stage of the transcription cycle. To investigate how the active compound(s) present in the crude extract of DEM40380 (DEM40380-CE) inhibit transcription, we first assessed if DEM40380-CE inhibited the synthesis of short abortive RNAs by WT *E. coli* RNAP on a linear

DNA template containing the T7A1 promoter. Specifically, we investigated inhibition of CpApU synthesis. In this instance, transcription was initiated with the dinucleotide primer CpA. It is well described Rif fails to inhibit synthesis of the first phosphodiester bond under such conditions, i.e. when transcription is initiated with a 5' non-phosphorylated dinucleotide. The principal mode of action of rifamycins is through steric occlusion of the translocating nascent transcript *following* formation of the first or second phosphodiester bond (Campbell et al., 2001, McClure and Cech, 1978). Interestingly, however, as shown in Figure 4.3, DEM40380-CE exhibited dose dependent inhibition of CpApU synthesis. This suggests DEM40380-CE targets a step of transcription initiation preceding formation of the first phosphodiester bond, or inhibits an element of the nucleotide addition cycle.

To establish if DEM40380-CE inhibits catalytic events, we assembled elongation complexes *in vitro* with fully complementary template and non-template strands, and 5'- radiolabelled RNA (Figure 4.3, B). By artificially assembling elongation complexes we circumvent transcription initiation events and can assess the effect of DEM40380-CE on transcription elongation. Indeed, as can be seen from figure 4.3, high resolution gel analysis of run-off RNA products synthesised by RNAP showed DEM40380-CE moderately inhibits transcription elongation. Thus, these data suggest the compound(s) present in DEM40380-CE can inhibit both initiation and elongation of transcription.
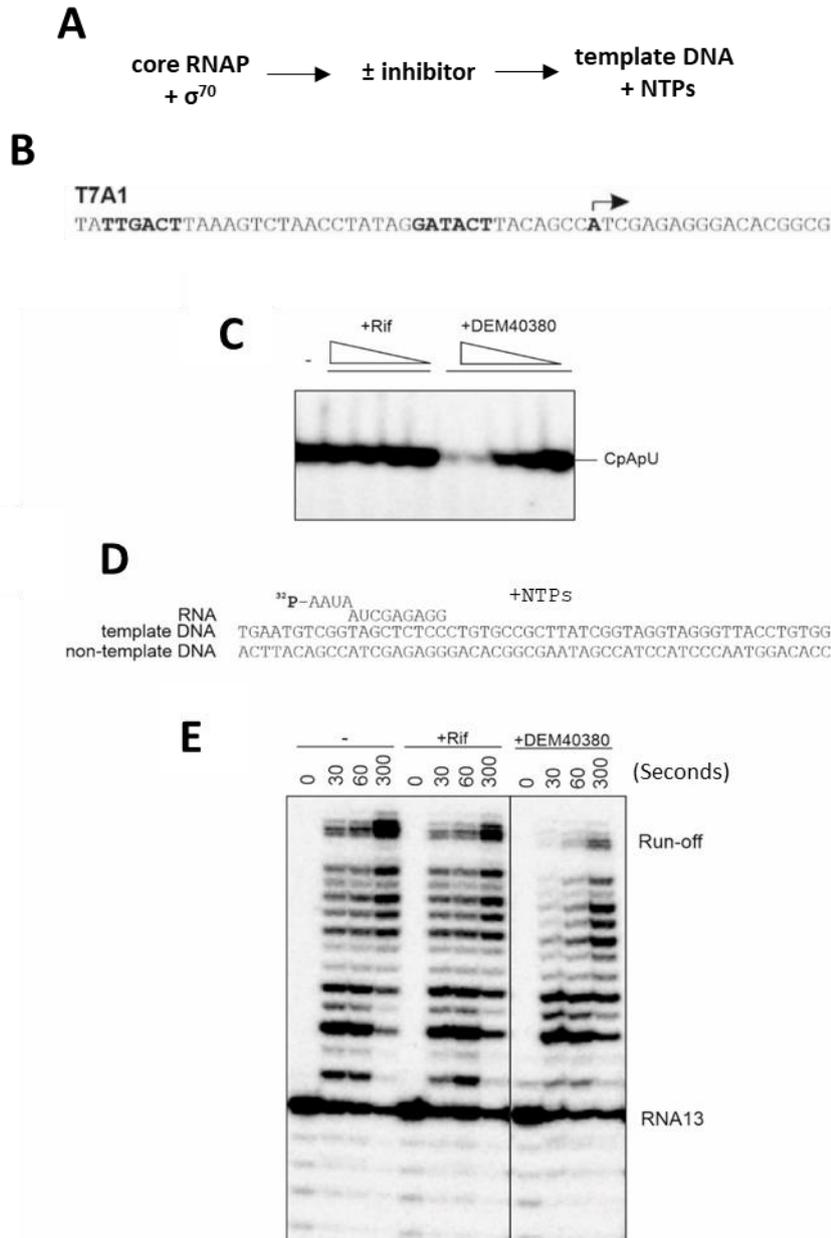
Figure 4.3 Extract of DEM40380 inhibits both transcription initiation and transcription elongation. (A) The reaction scheme of an *in vitro* transcription reaction from promoter DNA (B) The promoter DNA sequence of the T7A1 promoter used in C (C) Abortive synthesis of [^{32}P]-labelled CpApU by WT *E. coli* RNAP on linear DNA template containing the T7A1 promoter in the absence and presence of a serial dilution of Rif or a serial dilution of DEM40380 extract. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA. (D) The elongation complex (EC) used in C. Template DNA, non-template DNA, and 5'-[^{32}P]-labelled RNA are as indicated. (E) 5'-[^{32}P]-labelled RNA products synthesised from the EC scaffold shown in B, in the absence and presence of Rif (100µg/ml) (negative control) or DEM40380 extract (crude). 5'-[^{32}P]-RNA13 primer and 5'-[^{32}P]-RNA product are indicated.

4.2.3 DEM40380-CE contains a selective and a non-selective transcription inhibitor

To retain a high throughput nature to our initial screen, purification of bacterial extracts was kept minimal, comprising only a methanol extraction and single C18 solid phase extraction. However, to identify the exact compound(s) responsible for transcription inhibition, a more thorough purification was warranted. We therefore performed a second C18 Solid Phase Extraction in which the cartridge was eluted with incremental methanol (MeOH) elutions to fractionate DEM40380-CE. Fractions were evaporated to dryness and suspended in 50% MeOH. We subsequently analysed individual fractions in an *in vitro* transcription assay with WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. Interestingly, as can be seen in figure 4.4, two distinct fraction groups inhibited transcription by WT *E. coli* RNAP; the 20% elution and 70%-90% elutions, suggesting 2 inhibitory compounds may be produced by DEM40380. Moreover, the 70%-90% elutions inhibited transcription of full length RNA products with concurrent accumulation of short abortive products. This mechanism is typical of rifamycin inhibitors which inhibit transcription through steric hindrance of the translocating nascent transcript *following* synthesis of the first or second phosphodiester bond (McClure and Cech, 1978, Campbell et al., 2001).

To establish the selectivity of inhibitory fractions for bacterial RNAP, we assessed the effect of both the 20% elution fraction (DEM40380-F20) and the 80% elution fraction (DEM40380-F80) on *in vitro* transcription by RNA polymerase of the T7 bacteriophage (T7 RNAP) on a linear DNA template containing the T7 promoter sequence. T7 RNAP is a single subunit polymerase, structurally unrelated to multi-subunit RNAPs (Cheetham et al., 1999). Consequently, T7 RNAP can be used as a tool with which to identify non-specific inhibitors of transcription, such as the DNA binding compound echinomycin. Disappointingly, DEM40380-F20 inhibited transcription by T7 RNAP, suggesting the compound is a non-specific inhibitor of transcription. However, DEM40380-F80 doesn't inhibit T7 RNAP indicating the inhibitory compound(s) present in this particular fraction are selective for bacterial RNAP (Figure 4.4, B).
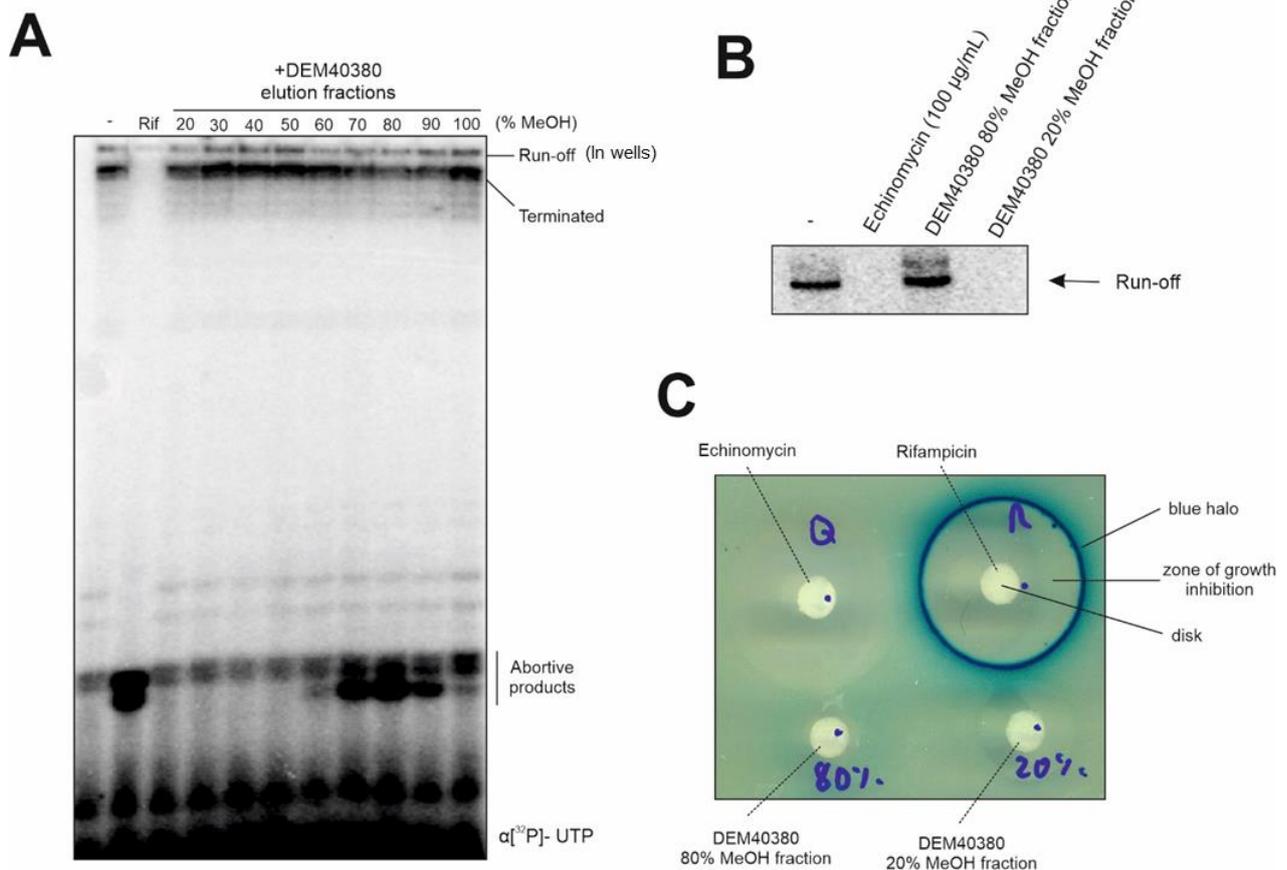
Figure 4.4 DEM40380 produces a specific and non-specific transcription inhibitor. (A) *in vitro* transcription by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of DEM40380 methanolic elution fractions. Fractions were obtained from incremental aqueous methanol elutions of DEM40380 crude extract loaded onto a C18 solid phase extraction cartridge. Fractions were subsequently dried and resuspended in 50% MeOH before in vitro transcription analysis. [^{32}P]-RNA products (Run-off and terminated) are indicated. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA. (B) *in vitro* transcription by RNAP of the T7 bacteriophage on a linear DNA template containing the T7 promoter, performed in the absence or presence of 80% and 20% DEM40380 methanolic elution fractions (attained as in A), or in the presence of DNA binding compound echinomycin (100µg/ml). 5'-[^{32}P]-RNA run-off product is shown. (C)   Disk diffusion assay with echinomycin, rifampicin, and 80% and 20% DEM40380 methanolic elution fractions (attained as in A). Paper disks soaked with antibiotic or respective fraction and were placed on LB agar plates infused with X-gal and lawn of yvgS *B. subtilis* reporter strain carrying the *lacZ* gene under the *HelD* promoter. This promoter is induced during partial inhibition of transcription (Hutter et al., 2004b). Note the blue halo at the frontier of the zone of growth inhibition in the case of transcription inhibitors.

To provide further clarification on the selectivity of our inhibitory fractions, we subjected DEM40380-F20 and DEM40380-F80 to a disk assay against the yvgS reporter strain. In such a disk assay, it is presumed specific inhibitors of RNAP produce a blue halo at the frontier of the zone of inhibition upon X-gal infused

agar plates. Like echinomycin, DEM40380-F20 fails to produce a blue halo at the frontier of the zone of growth inhibition, further substantiating the putative non-specific mechanism of transcription inhibition (Figure 4.4, C). However, DEM40380-F80, like rifampicin produces a visible blue halo at the frontier of the zone of growth inhibition, further corroborating the presence of an antibiotic targeting RNAP. Consequently, we selected DEM40380-F80 for further analysis.

4.2.4 DEM40380 produces the RNAP targeting compound Antibiotic A39079S-1

DEM40380-F80 inhibited transcription of full-length RNA transcripts (both terminated and run off) whilst concurrently causing an accumulation of short abortive products (Figure 4.4, A). This mechanism is typical of rifamycins, suggesting DEM40380-F80 may contain a 'rifamycin-like' compound. Consequently, we reanalysed DEM40380-F80 by LC-MS and performed a corresponding literature search in the dictionary of natural products. The HPLC chromatogram indicated a single prominent peak at 220nm absorbance, indicating sample homogeneity (Figure 4.5, A). Mass spectroscopic analysis of this peak showed 100% relative abundance of a species with m/z of 704.2779 Da. Moreover, a further species was identified with m/z of 1385.5495 Da corresponding to [2M+Na], where [M+Na] = 704.27. An additional species with m/z of 682.288 Da was identified and presumed to correspond to [M+H] (Figure 4.5, B). We therefore deduced DEM40380-F80 contains a compound with the exact mass of 681.288 Da. A search within the dictionary of natural products of 681.288 ± 0.01 Da yielded a single hit; 'Antibiotic A39079S-1', suggesting DEM40380-F80 contains this particular compound. Mass spectroscopic analysis of DEM40380 70 and 90% elution fractions showed they also contained Antibiotic A39079S-1. The structure of Antibiotic A39079S-1 is highly similar to the RNAP targeting ansamycin, rifamycin S (Rif-S), albeit with several subtle structural differences (Figure 4.5, C). The C27 methoxy group present in Rif-S is substituted by a hydroxyl group in Antibiotic A39079S-1. The methyl group at C16 present in Rif-S is absent in Antibiotic A39079S-1, and an additional methyl group is present at the C3 position of naphthoquinone moiety in Antibiotic A39079S-1.

Antibiotic A39079S-1 was first isolated from *Streptomyces spheroides* NRRL 15600 in Vancouver, Canada (Boeck, 1985). It is a broad spectrum antibiotic with a previously uncharacterised mechanism of action. Our results thus far and structural similarity to Rif-S indicate Antibiotic A39079S-1 targets RNAP through a mechanism of steric occlusion in which the compound binds within the Rif pocket and sterically blocks translocation of the nascent transcript. To further assess if Antibiotic A39079S-1 targets the Rif pocket on RNAP, we reassessed DEM40380-F80 in an *in vitro* transcription assay with Rif^R *E. coli* RNAP bearing βH526Y on a linear DNA template containing the T7A1 promoter. DEM40380-F80 was unable to inhibit transcription performed by the mutant RNAP, suggesting Antibiotic A39079S-1 targets RNAP at the Rif binding pocket (Figure 4.5, D). Previous inhibition of βH526Y RNAP by DEM40380-CE was presumably due to the presence of another inhibitory compound within the crude fraction.
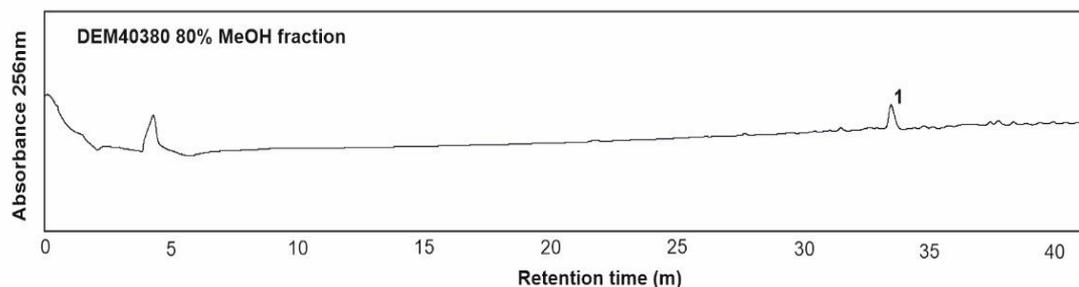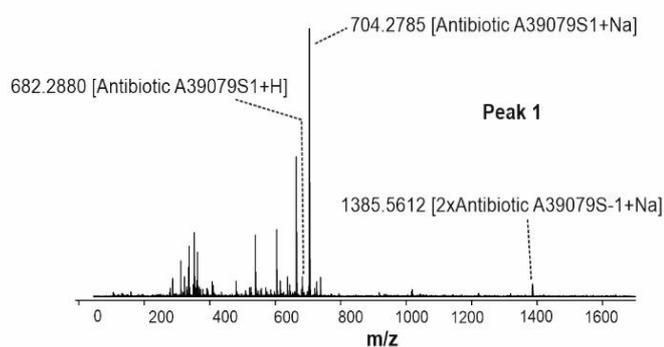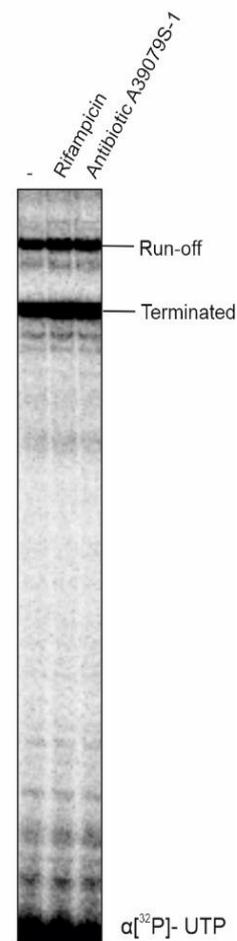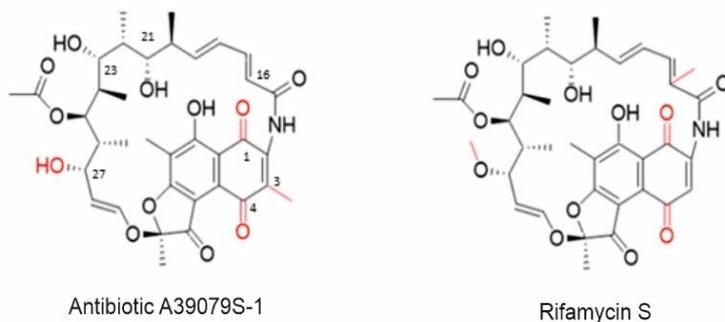
Figure 4.5 DEM40380 produces the rifamycin class compound 'Antibiotic A39079S-1' (A) A representative reverse phase HPLC trace of DEM40380 80% MeOH fraction (DEM40380-F80) at 256nm absorbance. (B) Positive ion mass spectrum corresponding to peak '1' identified in A. Probable Ion adducts of 'Antibiotic A39079S-1' are indicated. (C) Structural formulae of Antibiotic A39079S-1 (left) and Rifamycin S (right). Structural differences are indicated in red. (D) *in vitro* transcription by Rif[R] βH526Y RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of rifampicin or antibiotic A39079S-1. [32P]-RNA products (Run-off and terminated) are indicated. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA.

4.5 Discussion

The growing prevalence of infections resistant to antibiotics demands compounds with novel modes of action and novel binding interactions. In this chapter, we aimed to identify novel RNAP inhibitors produced by actinomycete bacteria from a shortlisted library of actinomycetes compiled by our industrial collaborators at Demuris. Actinomycetes form a rich reservoir of active natural products, from which many of our current antibiotics are derived directly, or synthesised from. Demuris possess an extensive collection of actinomycete bacteria from a variety of marine and terrestrial sources. In an effort to identify novel transcription targeting antibiotics, Demuris performed a screen of their collection in which strains were screened against the *Bacillus subtilus yvgS* reporter strain which has *lacZ* fused to the promoter of the bacterial helicase *HelD*. This particular promoter is upregulated in response to sub-inhibitory concentrations of the transcription inhibitor Rif. Consequently, it was assumed this particular strain reports on producers of specific transcription inhibitors. However, our work here has shown most, if not all, strains forwarded for investigation produce rifamycin class antibiotics. This indicates producers of rifamycins may be far more prevalent than producers of other transcription inhibitors, or the yvgS reporter strain fails to report transcription inhibitors other than rifamycins. Indeed, work at Demuris following the course of this project showed certain actinomycete derived transcription inhibitors, such as streptolydigin, fail to activate *yvgS* (unpublished), illustrating the limited suitability of the yvgS reporter in drug discovery. Consequently, a reporter strain with broader scope for activation is required if actinomycete bacteria are to be comprehensively examined for novel RNAP inhibitors.

Nevertheless, *Streptomyces* DEM40380, isolated from the Atacama Desert in Chile, was identified as a likely producer of a previously unknown RNAP inhibitor. *In vitro* transcription assays combined with mass spectroscopic analysis identified DEM40380 as a producer of Antibiotic A39079S-1, a broad spectrum ansamycin antibiotic with a previously undefined mechanism of action. Here, we have shown Antibiotic A39079S-1 exhibited a mechanism of action typical of rifamycin inhibitors, in which the compound inhibits transcription of full length RNA products whilst concurrently accumulating short abortive RNAs. This indicates the

compound acts through steric occlusion of the extending nascent transcript following initial phosphodiester bond synthesis (McClure and Cech, 1978, Campbell et al., 2001). Furthermore, the antibiotic lost activity at RNAP bearing the Rif[R] mutation βH526Y, indicating the compound targets the Rif binding pocket on RNAP (Campbell et al., 2001). Indeed, Antibiotic A39079S-1 has a highly similar structure to the natural product rifamycin S, a potent inhibitor of RNAP (Sensi et al., 1960). The compound possesses the distinctive naphthelenic, 17-mer *ansa* chain structure characteristic of rifamycin antibiotics. Yet, the activity of antibiotic A39079S-1 sheds new light on the structure-activity relationship of rifamycins. Indeed, in the context of antibiotic A39079S-1, the methoxy group at C27, common to most rifamycins, can seemingly be cleaved to a hydroxyl without abolishing activity. Furthermore, the absence of a C16 methyl group in Antibiotic A39079S-1 suggests this substituent is not required for the correct conformation of essential oxygen functionalities at C1, C8, C21 and C23. This observation may point to a site on the *ansa-* bridge where small substituents can be introduced successfully to the compound.

To summarise, we have shown the previously uncharacterised compound Antibiotic A39079S-1 mediates antibiotic activity through selective inhibition of bacterial RNAP. The compound likely binds within the Rif binding pocket on RNAP to block the extension of the nascent transcript following initial phosphodiester bond formation. We have also shown actinomycetes remain an abundant source of active natural products. However, the methods by which we screen for producers of these compounds requires further work if we are to prevent re-discovery of non-novel compound classes. Indeed, our identification of strain DEM4038 as a producer of Antibiotic A39079S-1 provides an example of the difficulties in identifying truly novel compounds in the natural product drug discovery process.

Chapter 5. Ureidothiophene inhibits recognition of -10 promotor element by targeting regulatory region 1.2 of sigma subunit

5.1 Introduction

The synthetic antibiotic ureidothiophene (Urd) (Figure 5.1, A) was discovered in a high-throughput screen of some 250,000 commercially available compounds in which activity against *S. aureus* RNAP holoenzyme was assessed *in* vitro (Arhin et al., 2006). The compound was shown to be highly active against *S. aureus* RNAP *in vitro*, with an IC_{50} of ~1µM. Urd possessed a narrow spectrum of activity against *S. aureus* ATCC 13709 and *S. epidermidis* with a ~MIC of 1µg/ml and 0.25µg/ml, respectively (Arhin et al., 2006). An isopropyl derivative of Urd was shown to inhibit RNA and protein synthesis, but not DNA synthesis by *S. aureus* strain RN4220; an effect typical of selective RNAP inhibition. Additionally, the compound retained activity against Rif resistant strains of *S. aureus* suggesting the binding site of Urd is different to that of Rif (Arhin et al., 2006). However, the exact binding site and mechanism of inhibition by Urd at RNAP remained unknown. This study aimed to elucidate the molecular mechanism by which Urd inhibits bacterial RNAP.

5.2 Results

5.2.1 Ureidothiophene inhibits RNA polymerases *in vitro*

Firstly, we assessed the effects of Urd on *in vitro* transcription by the wild-type *E. coli* RNAP, the most extensively characterised bacterial RNAP. Urd inhibited transcription on a linear DNA template containing *lac*UV5 promoter (IC_{50} ~15.1 ± 8.1µg/ml) (Figure 5.1, B). A decrease in full length transcript synthesis coincided with a corresponding decrease in the synthesis of short abortive products. This mechanism is seemingly different to that of Rif which typically causes an accumulation of short abortive products whilst inhibiting full length transcript synthesis (McClure and Cech, 1978, Campbell et al., 2001). Indeed, Urd inhibits synthesis of both the tri- and tetra- nucleotide abortive products ApApU and ApApUpU in an abortive transcription assay (IC_{50} ~18.5 ± 1.9 µg/ml) (Figure 5.1, C).
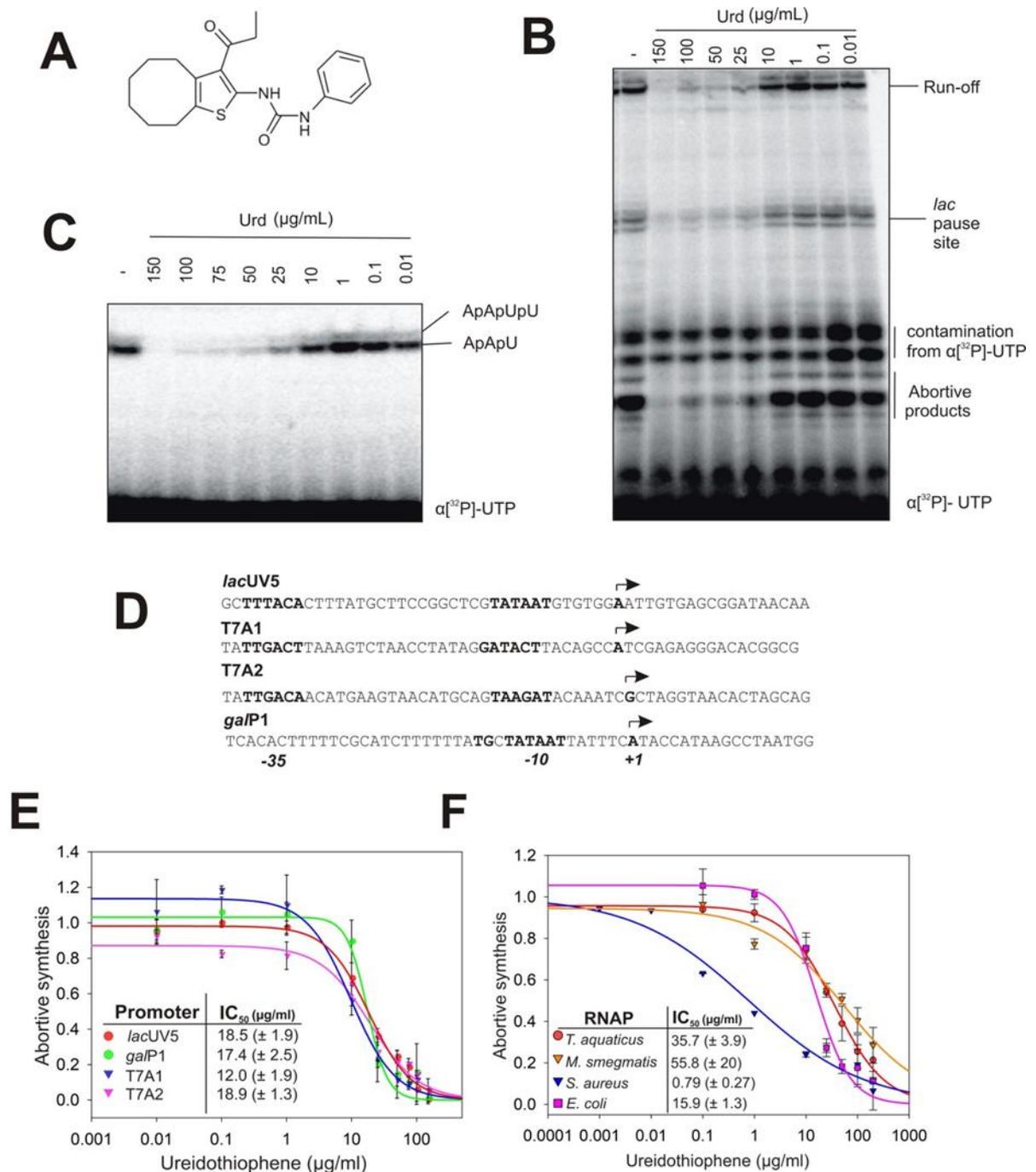
Figure 5.1 Ureidothiophene (Urd) inhibits bacterial RNA polymerases. (A) Structural formula of Urd. (B) Urd inhibition of *in vitro* transcription performed by WT *E. coli* RNAP on a linear DNA template containing the *lac*UV5 promoter. [32P]-RNA products (Run-off and terminated) are indicated. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer ApA. (C) Urd inhibition of abortive synthesis of [32P]-labelled ApApU and ApApUpU by WT *E. coli* RNAP holoenzyme on linear DNA template containing the *lacUV5* promoter. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer ApA. (D) Sequence of promoters used in E. Promoter elements and TSS are indicated. (E) Urd inhibition of *in vitro* transcription by *E. coli* RNAP holoenzyme on linear DNA templates containing different promoter sequences as indicated in D. Error bars are ± SD from 3 independent experiments. (F) Urd inhibition of *in vitro* transcription by bacterial RNAP holoenzymes from different bacteria on linear DNA templates containing the T7A1 promoter. Error bars are ± SD from at least 3 independent experiments.

73

To ensure the transcription inhibition by Urd wasn't specific to the *lac*UV5 promoter, we assessed Urd inhibition of abortive transcription by wild-type *E. coli* RNAP on linear DNA templates containing the T7A1, T7A2 promoter and the extended -10 *gal*P1 promoter (Figure 5.1, D). These promoters represent differing deviations from consensus -10 and -35 promoter element sequences. For example, the *gal*P1 promoter sequence possesses a 5'-TG-3' motif 2 bps upstream of the -10 element but does not possess a recognisable -35 element. On the other hand, T7A1 and T7A2 possess distinct -10 and -35 promoter element sequences whilst also possessing differing lengths within the spacer region (see section 1.2.4 for further details on promoter architecture). Urd inhibited abortive transcription on T7A1, T7A2 and the extended -10 *gal*P1 promoters with comparable potency (Figure 5.1, E). This suggests Urd inhibition is not sequence specific *per se*, and possesses a general mechanism of inhibition. We also assessed the ability of Urd to inhibit transcription by different bacterial RNAPs (Figure 5.1, F). In this experiment we used the T7A1 promoter as template. Consistent with previous observations, *S. aureus* RNAP was highly susceptible to Urd with an IC_{50} ~ 0.3 (± 0.3) µg/ml (Arhin et al., 2006). In contrast, *T. thermophilus* and *M. smegmatis* RNAPs were much less sensitive to Urd.

5.2.2 Ureidothiophene is an inhibitor of transcription initiation

Concurrent inhibition of both abortive and run-off transcription suggests Urd may inhibit nucleotide binding or catalysis. Therefore, we analysed the ability of Urd to inhibit single and multiple nucleotide addition by elongation complexes formed by *E. coli* RNAP core. Elongation complexes were assembled with synthetic oligonucleotides wherein the elongation complex contained fully complementary template and non-template strands and 5′- radiolabelled RNA (Figure 5.2, A). As seen from Figure 5.2, even a high concentration (100 µg/mL) of Urd had no effect on single nucleotide and multiple nucleotide RNA extension, indicating that the inhibitor doesn't effect NTP binding or catalysis. This suggests Urd targets an early stage of transcription initiation.
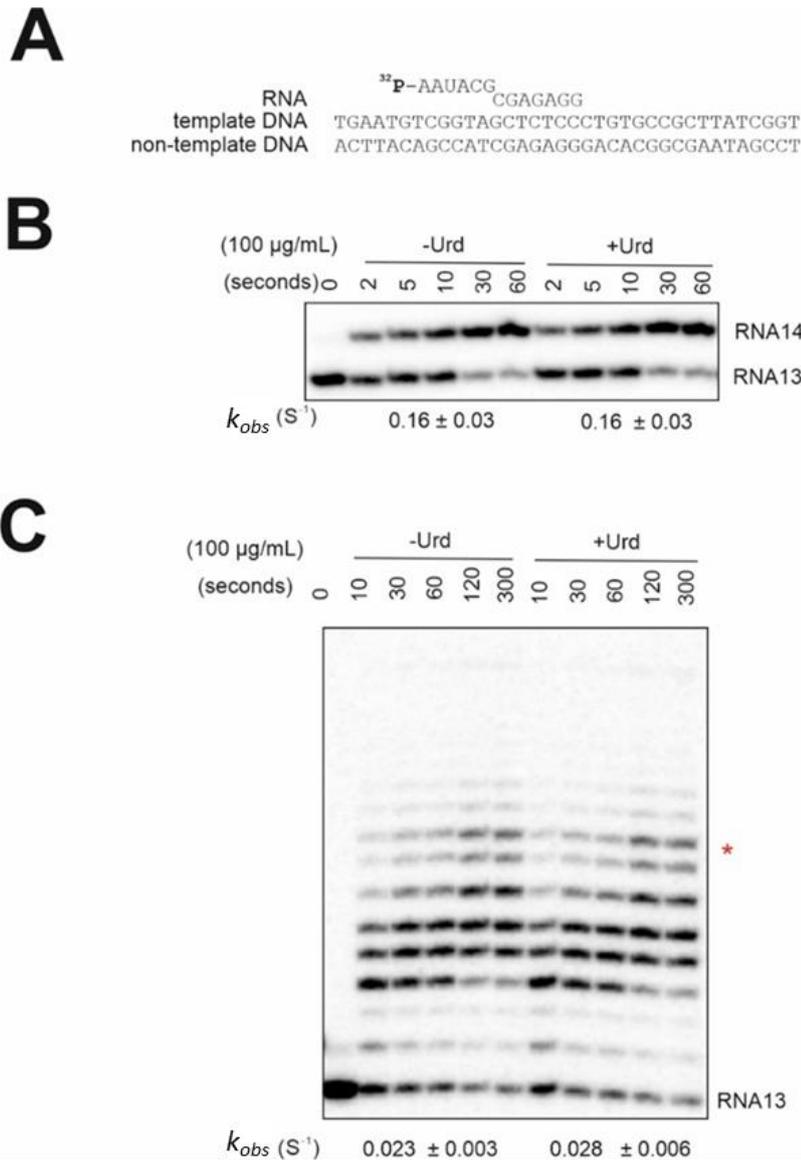
Figure 5.2. Ureidothiophene does not inhibit transcription elongation. (A) The elongation complex scaffold (EC) used in (B) and (C). Template DNA, non-template DNA, and 5'-[^{32}P]-labelled RNA are as indicated. (B) 5'-[^{32}P]-labelled RNA products synthesised from single nucleotide addition of the EC scaffold shown in A, in the absence and presence of Urd (100µg/ml). 5'-[^{32}P]-RNA13 primer and 5'-[^{32}P]-RNA14 product are indicated. Rate constants are shown below the gels (numbers that follow the ± sign are standard errors). (C) 5'-[^{32}P]-labelled RNA products synthesised from multiple nucleotide addition of the EC scaffold shown in A, in the absence and presence of Urd (100µg/ml). 5'-[^{32}P]-RNA13 primer indicated, and quantified 5'-[^{32}P]-RNA product is marked by asterisk. Rate constants are shown below the gels (numbers that follow the ± sign are standard errors)
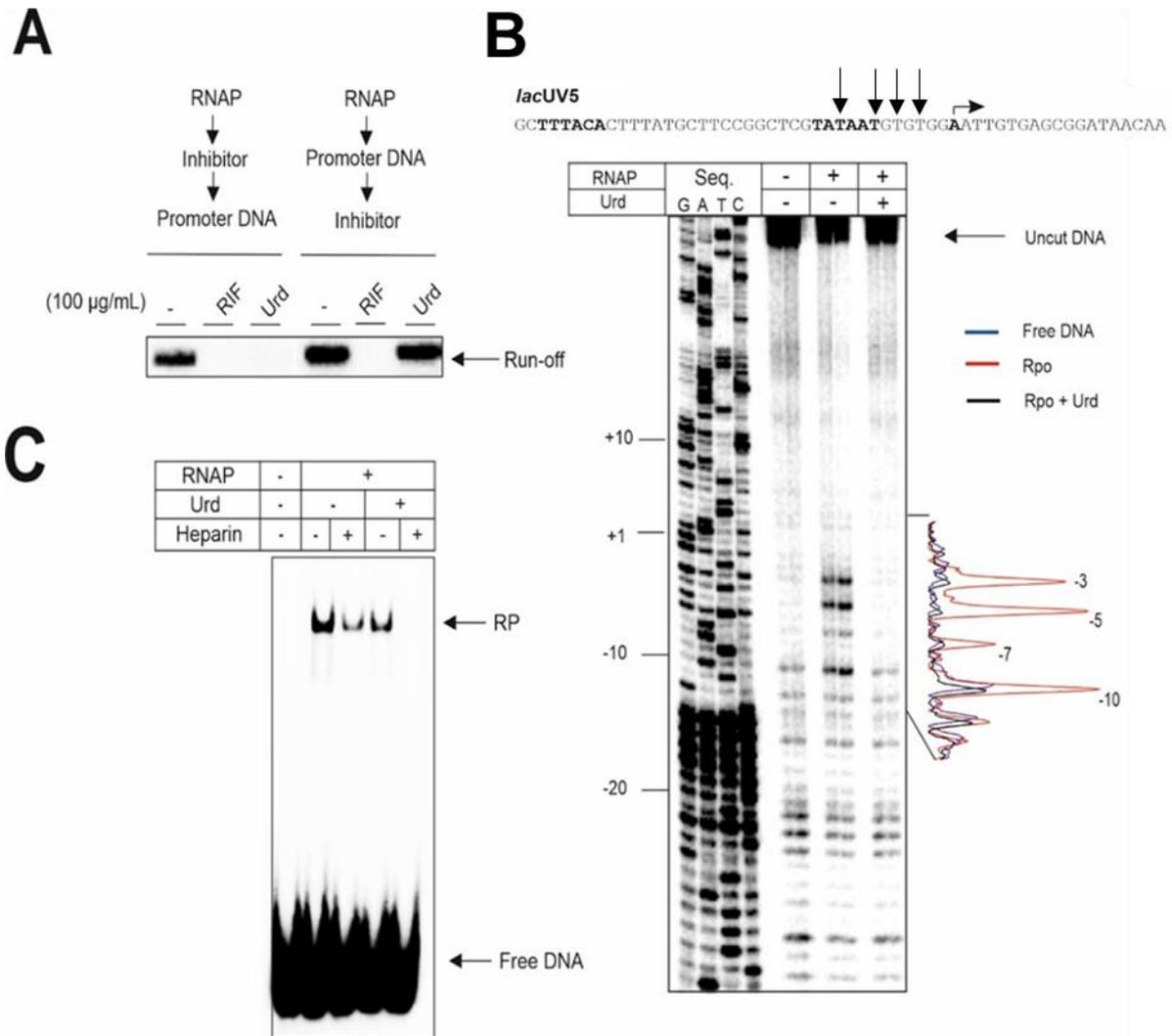
Figure 5.3 Urd inhibits formation of the RNAP open promoter complex. (A) Effect of order of addition of Urd, prior to, or following addition of promoter DNA, on inhibition of *in vitro* transcription by *E. coli* RNAP holoenzyme. (B) KMnO$_4$ probing of RNAP-*lac*UV5 promoter complexes assembled in the absence and presence of Urd (100µg/ml) Non-template strand was 5'-[^{32}P]-labelled. The *lac*UV5 promoter sequence is shown above the gel with sensitive thymines indicated by black arrows. Profiles to the right of the gel are representative scans for free DNA (blue), RPo (red) and RPo + Urd (black). Position of thymine's susceptible to modification by KMnO$_4$ in RPo are indicated. Sanger sequencing of the promoter fragment is shown on the left hand portion of the gel. (C) Electrophoretic mobility shift assay (EMSA) of RNAP-*lac*UV5 promoter complexes assembled in the absence and presence of Urd (100µg/ml). Complexes were further challenged with heparin as indicated. Non-template strand was 5'-[^{32}P]-labelled. Complexes were resolved in native 4.5% PAGE. The position of the RNAP-*lac*UV5 promoter complexes (RP) and non-bound DNA (Free DNA) are shown.
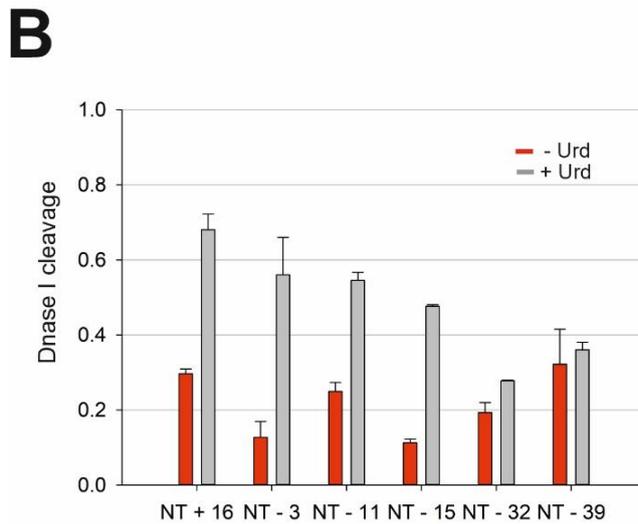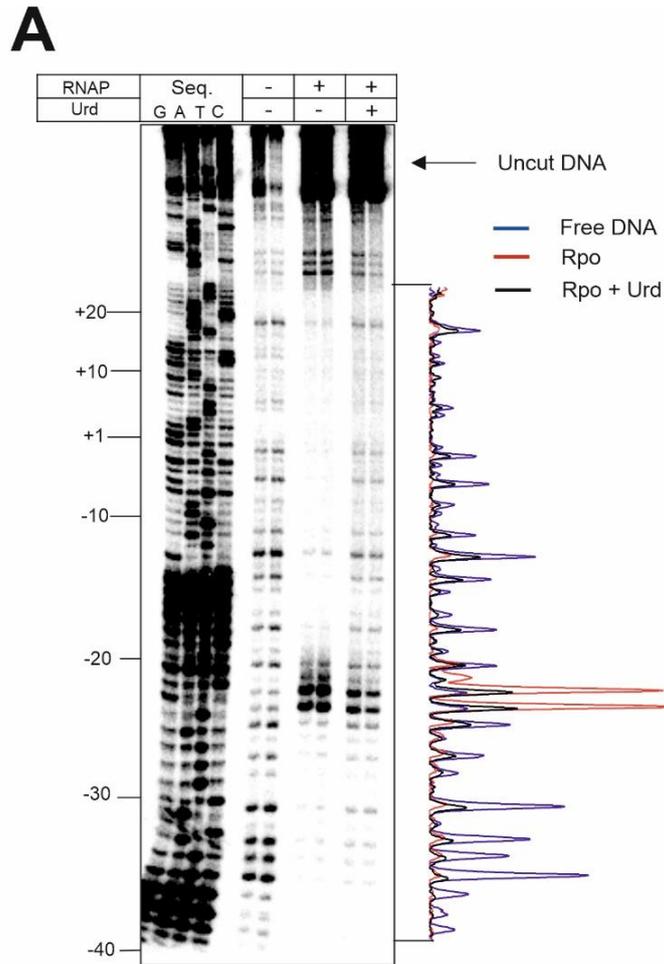
Figure 5.4 Urd prevents recognition of downstream promoter DNA. (A) DNAase I probing of RNAP-*lac*UV5 promoter complexes assembled in the absence and presence of Urd (100µg/ml) Non-template strand was 5'-[³²P]-labelled. Profiles to the right of the gel are representative scans for free DNA (blue), RPo (red) and RPo + Urd (black). Sanger sequencing of the promoter fragment is shown on the left hand portion of the gel. (B) Bar graph showing quantification of

DNase I footprinting data shown in A. Chart indicates peak area values for indicated bands within promoter DNA fragment in RPo in the presence and absence of Urd, normalised to peak area values for corresponding position in free DNA. Error bars are ± SD.

5.2.3 Ureidothiophene prevents RNAP interaction with downstream promoter DNA

We found Urd inhibition possesses a marked order-of-addition dependency. The compound inhibits initiation of transcription at a stage prior to formation of promoter open complex as it does not have any effect on transcription when added after formation of RPo (Figure 5.3, A). We therefore, analysed if Urd targets formation of the RPo by $KMnO_4$ footprinting, which probes unpaired thymidine residues in the melted region of the RPo. Linear DNA fragment containing the *lac*UV5 promoter was radiolabelled at the 5' end of the non-template strand. As can be seen from Figure 5.3, B, Urd (100µg/ml) added before mixing RNAP and promoter DNA completely inhibited formation of RPo; thymidines in positions -10, -7, -5 and -3 that were melted in RPo, remained double-stranded in the presence of Urd.

Urd may block interactions of RNAP with DNA or prevent crucial interactions of RNAP with promoter DNA that precede the melting and/or loading of promoter DNA in to the RNAP active-site cleft. In order to distinguish between these possibilities, we analysed RNAP-promoter formation by electrophoretic mobility shift assay (EMSA) (Straney and Crothers, 1985). *E. coli* RNAP was treated with Urd (100 µg/ml) and then incubated with radiolabelled *lac*UV5 promoter DNA. As can be seen from Figure 5.3, C, Urd did not abolish interaction of RNAP with promoter DNA. However, a challenge with polyanion heparin leads to destruction of the complexes formed in the presence of Urd. Heparin sequesters free RNAP and is also thought destabilise RPc, however has less effect on stable RPo (Walter et al., 1967). We therefore conclude that Urd blocks a stage of isomerisation into RPo, but does not abolish recognition of promoter DNA by RNAP.

To understand the nature of Urd/RNAP/promoter complexes, we performed DNase I footprinting of promoter complexes in the absence or presence of Urd (100µg/ml) added before mixing RNAP and the promoter DNA (*lac*UV5 with 5'-

radiolabelled non-template strand). DNase I is a non-specific endonuclease that can be used to identify DNA regions protected by DNA binding proteins such as RNAP. Our results indicate Urd doesn't cause a significant change in protection in the upstream promoter regions from positions -39 to -25 (Figure 5.4), suggesting that σR4.2 is engaged with the -35 promoter element in the presence of Urd. However, a large difference in protection pattern is observed downstream of the -35 promoter region. Hypersensitive sites at positions -23 and -24 on the non-template strand, that arise from distortion of the 18 base pair spacer region between the -10 and -35 promoter elements (Carpousis and Gralla, 1985), have diminished sensitivity to DNAse I digestion in the presence of Urd (Figure 5.2, A), suggesting the -10 may be disengaged by σR2 . Indeed, Urd causes a strong deprotection of nearly all bases downstream of position -20 up to +18. Notably, -11 adenosine residue, essential for recognition of -10 element (Bae et al., 2015a, Roberts and Roberts, 1996), is deprotected in the presence of Urd indicating the -10 element is unable to form stable contacts with σR2 (Figure 5.4, A, B). We therefore conclude that Urd doesn't inhibit binding of the -35 promoter element, however the inhibitor prevents melting of -10 element and/or prevents correct loading of DNA within the active site cleft and DNA binding to the β/ β' channel.

5.2.4 Ureidothiophene doesn't inhibit binding of DNA to downstream DNA-binding channel or loading of template DNA into the active cleft

Urd may occlude the access of single stranded template DNA into the active site cleft and/or occlude the binding of duplex DNA into the β and β' downstream DNA binding channel and/or affect the recognition of downstream promoter elements essential for formation of RPo. Two previously described inhibitors of RNAP, fidaxomicin (Fdx) and ripostatin (Rip), were shown to inhibit isomerisation into RPo. Fdx blocks RPo formation by locking the RNAP clamp in an open conformation by binding to the switch region, the molecular hinge that facilitates clamp movement, and by blocking binding of template DNA within the active centre (Boyaci et al., 2018, Lin et al., 2017a, Tupin et al., 2010a). Like Urd, Fdx doesn't affect the binding of upstream promoter elements whilst destabilising binding of downstream promoter DNA (Tupin et al., 2010a, Morichaud et al., 2016). Rip also binds to the switch region but inhibits isomerisation to RPo at a later stage than Fdx by trapping a promoter complex with a partially melted

transcription bubble that fails to propagate to the transcription start site. (Mukhopadhyay et al., 2008).

We analysed if the mechanism of action of Urd has similarities with that of Fdx and Rip by using single-stranded promoter of the origin of replication of M13 bacteriophage (M13ori; Figure 5.5, A) (Bochkareva and Zenkin, 2013, Zenkin and Severinov, 2004, Zenkin et al., 2006). M13ori forms a hairpin which is recognised by the downstream DNA-binding channel of RNAP but does not require -10 or -35 elements for its function. Neither DNA upstream of the +4 position of the 'non-template' strand, nor upstream of -3 position of the 'template' stand, are required for binding of the M13 hairpin by RNAP. Although, σ^{70} is required for the initiation
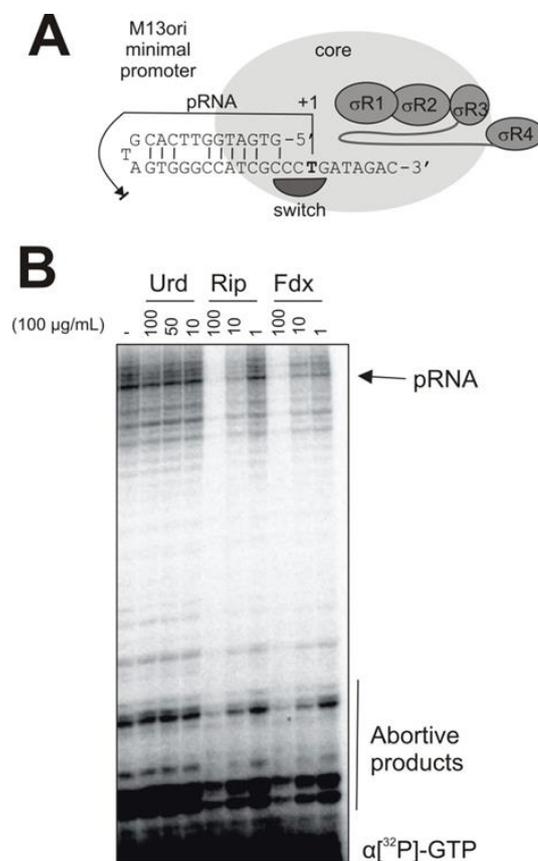
Figure 5.5 Ureidothiophene doesn't inhibit binding of DNA to downstream DNA-binding channel or loading of template DNA into the active cleft (A) Structure of the single-stranded M13 minimal promoter recognized by downstream-DNA-binding channel of RNAP, and which binding does not depend on -10/-35 elements or on the σ^{70} subunit. 18nt RNA product (pRNA) synthesized on M13ori promoter is shown with an arrow. (B) *in vitro* transcription of an 18nt pRNA primer performed by WT *E. coli* RNAP holoenzyme, on single stranded M13 minimal promoter template shown in A. Experiment was performed in the absence and presence of ureidothiophene, and switch region targeting inhibitors ripostatin and fidaxomicin. [32P]-pRNA and [32P]-abortive products are indicated.

of transcription at M13ori, σ^{70} is not required for the binding of M13ori to RNAP (Zenkin et al., 2006, Bochkareva and Zenkin, 2013, Zenkin and Severinov, 2004). Thus, M13ori can be used to assess if a compound inhibits downstream DNA-binding and/or loading of the single-stranded DNA into the active cleft (Zenkin et al., 2006, Zenkin and Severinov, 2004). Therefore, we tested transcription by *E. coli* holoenzyme on minimal M13ori that leads to the formation of an 18nt primer RNA (pRNA; Figure 5.5, A, B), in the presence of the inhibitors Urd, Fdx and Rip. As expected, Fdx and Rip strongly inhibit transcription on M13ori (Fig 5B) (Srivastava et al., 2011, Tupin et al., 2010a). However, Urd has no effect of formation of pRNA. Results indicate that, unlike Fdx and Rip (Tupin et al., 2010a, Mukhopadhyay et al., 2008), Urd does not occlude the binding of single stranded template DNA into the active site and does not prevent binding of downstream DNA duplex to the downstream DNA channel. Taken together, the above data suggest that Urd inhibits isomerisation from RPc to RPo by precluding melting of the -10 promoter element, potentially through a previously unseen mode of action.

5.2.5 Ureidothiophene targets σR1.2

To delineate the putative binding site of Urd, we isolated an *S. epidermidis* spontaneous Urd resistant mutant conferring resistance to Urd termed *Urd-Mut1* (Figure 5.6, A). *Urd-Mut1* was isolated from LB agar plates containing 12.5 µg/ml Urd (4 x MIC) and 2.5% Pluronic F68. The Pluronic F68 copolymer was used in this instance to improve solubility of Urd, and is further known to improve penetration of small molecules across cellular membranes. Ureidothiophene resistant mutants were identified by their growth on this media and confirmed by re-streaking on the same media, leading to the identification of a single mutant - *Urd-Mut1*. Note that the low number of mutants identified may perhaps be due to a low resistance frequency in the *S. epidermidis* strain used here and/or insufficient plating of cells onto the growth media (see section 2.15). Genome sequencing revealed a sole amino-acid substitution in the genome of the *Urd-Mut1*; an E105Q mutation in the *rpoD* gene encoding the primary sigma factor, SigA. The mutation is located within σR1.2 (Figure 5.6, B). To confirm this mutation is responsible for

A

Organism/isolate	Strain Reference	Urd MIC (µg/ml)
Staphlyoccous epidermidis	*ATCC 12228*	3.125
Staphlyoccous epidermidis - Mut 1	*ATCC 12228* with *rpoD* E105Q mutation	>100

B

90°

C

S.epidermidis WT σ^A holoenzyme   *S.epidermidis* E105Q σ^A holoenzyme

Urd (100 µg/mL)   −   +   −   +

←CpApU

Inhbition (%)   −   75.1 (± 1.4)   −   28.0 (± 7.1)

E

RNAP	IC$_{50}$ (µg/ml)
WT	15.12 (± 8.05)
Sigma^{70} E104Q	90.81 (± 10.41)

Transcription (y-axis)

Ureidothiophene (µg/ml) (x-axis)

F

σ^{70} E381

β' I290

σ^{70} E104

NT -7
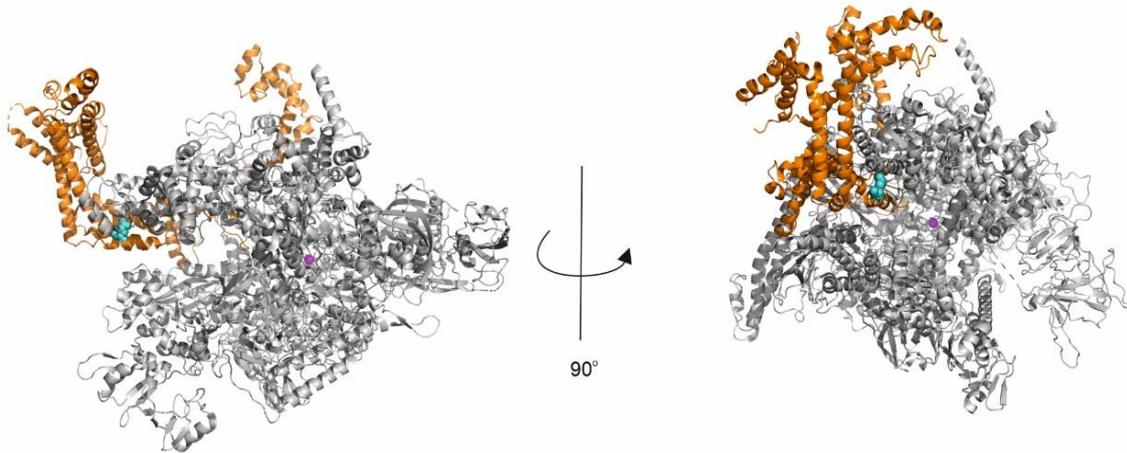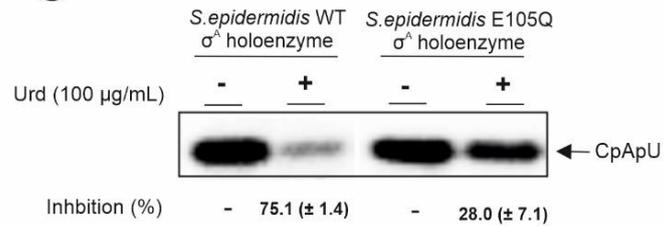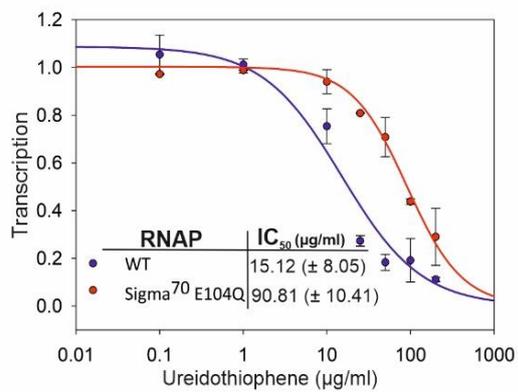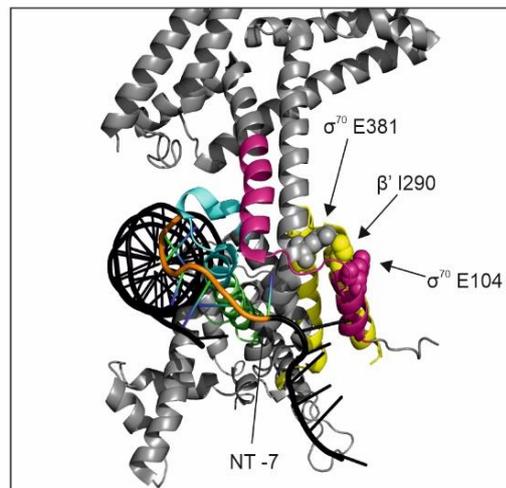
Figure 5.6 Urd inhibits RNAP by targeting σR1.2. (A) Table indicating the MIC values of *S. epidermidis* ATCC12228 and *Urd-Mut1*, a spontaneous Urd resistant *S. epidermidis* ATCC12228 with an σ^A E105Q mutation. MIC was gathered by solid agar dilution method (Wiegand et al., 2008). (B) Orthogonal views of the location of *S. epidermidis* σ^A E105 residue (E104; *E. coli* numbering) mapped onto the structure of *E. coli* holoenzyme. Core RNAP is depicted in grey ribbon model. σ^{70} is depicted in orange with residue E104 indicated depicted as blue sphere model. (C) Abortive synthesis of [^{32}P]-labelled CpApU by cellular WT *S. epidermidis* RNAP and cellular σ^A E105 *S. epidermidis* RNAP on linear DNA template containing the T7A1 promoter in the absence and presence of Urd. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA. % inhibition by Urd normalised. (E) Urd inhibition of *in vitro* transcription by WT *E. coli* RNAP holoenzyme and σ^{70} E104Q *E. coli* RNAP holoenzyme. Error bars are ± SD from 3 independent experiments. (F) Structural recognition of the promoter -10 element by σR2, and proposed allosteric switch of σR1.2 within WT *E. coli* holoenzyme. Promoter DNA is depicted in black, with the -10 element depicted with atomic colouring, non-template base at -7 position is indicated. σ^{70} is depicted in gray, and the RNAP β' coiled-coil domain depicted in yellow. Residues within the proposed σR1.2 allosteric switch are indicated in sphere model. σR1.2 (residues 96-126) is depicted in pink, σR2.3 (residues 416 to 434) is depicted cyan, and 2.4 (residues 435-452) is coloured green (Feklistov and Darst, 2011).

the resistance phenotype, we purified cellular RNAP holoenzyme from wild-type *S. epidermidis* and *Urd-Mut1,* and analysed sensitivity to Urd on the T7A1 promoter. Unlike the wild-type *S. epidermidis* RNAP, holoenzyme bearing the E105Q mutation is highly resistant to inhibition by Urd (Figure 5.6, C). This result confirms RNAP is a cellular target of Urd and also indicates the E105Q mutation underlies the observed resistance phenotype. To further corroborate this finding, we introduced the corresponding mutation into the *E. coli* σ^{70} subunit (E104Q; *E. coli*) and assessed the effect of the mutation on *E.coli* holoenzyme sensitivity to Urd.

Indeed, the mutant holoenzyme was almost 6-fold more resistant to Urd than the WT RNAP (Figure 5.6, E). σR1.2 is implicated in the formation of RPo (Haugen et al., 2006, Zenkin et al., 2007, Wilson and Dombroski, 1997, Baldwin and Dombroski, 2001). Our results therefore suggest that inhibition of RPo formation by Urd may involve σR1.2.

5.4 Discussion

In this study, we have shown the transcription inhibitor Urd inhibits initiation of transcription by preventing the formation of a transcription competent open promoter complex. Our data indicate Urd inhibition of RPo formation may involve σR1.2. This particular σ sub-region is implicated in the formation of open promoter complexes by regulating recognition of downstream promoter elements

(Baldwin and Dombroski, 2001, Zenkin et al., 2007). Indeed, here, we identify a mechanism of transcription inhibition involving σR1.2 wherein Urd prevents melting of the -10 promoter element to prevent formation of an open promoter complex.

Our data show Urd, like other inhibitors of open complex formation, acts at the promoter melting stage of transcription initiation. However, we have shown Urd acts through a different mechanism to previously described inhibitors of RPo formation, as the compound fails to inhibit loading of template DNA within the active centre. Fdx prevents recognition of the promoter -10 element and directly targets loading of single stranded template DNA into the active centre cleft, a detail supported by our own findings at M13ori DNA template (Tupin et al., 2010a). Urd, however, despite also preventing simultaneous recognition of the -10 and -35 elements, appears not to impede the interaction of single stranded template DNA at the active site, showing that Urd doesn't inhibit RPo formation by targeting template loading. This is highly indicative Urd acts through a mechanism different to that of other inhibitors of RPo formation.

A series of 2nd generation ureidothiophene-2-carboxylic acids (2G-Urds) structurally related to Urd were shown to possess mutually exclusive binding with Myx, Rip and Cor; switch region targeting ligands that bind to and lock the RNAP cleft in a *closed* conformation (Mukhopadhyay et al., 2008, Belogurov et al., 2008, Fruth et al., 2014). Fruth *et al* used saturation transfer difference NMR to show 2G-Urds competitively bind with these closed cleft locking ligands (Fruth et al., 2014). In this case, the authors concluded that 2G-Urds competitively bind at the same site on RNAP - the switch region. However, these aforementioned inhibitors target a later promoter melting step in which the promoter DNA is partially melted but propagation to the transcription start site is obstructed due to clamp conformation (Boyaci et al., 2019). This is dissimilar to the earlier stage of promoter complex formation targeted by Urd in which complexes possess an entirely closed transcription bubble. If Urd binds at the same site as 2nd generation analogues, we propose an alternative explanation. It is tempting to speculate this competitive binding does not arise from direct competition at the switch region binding site but instead arises from altered clamp conformation induced by ureidothiophenes that precludes the closed clamp conformation likely

required for Myx, Cor and Rip binding. Note that Myx, Cor and Rip are *closed* cleft locking ligands and therefore most probably require a closed clamp conformation in order for the compound to bind RNAP. Indeed, biochemical data from the same study suggests 2G-Urd compounds do not interact with RNAP residues crucial for Myx binding. We therefore suggest that Urd may affect conformation of the RNAP clamp, in addition to targeting σR1.2, predisposing an open or partly closed clamp conformation that may influence recognition of the -10 and -35 elements, but not through a means identical to Fdx in which binding of template DNA within the active centre is affected.

We have shown that Urd transcription inhibition is dependent on σ factor sub-region 1.2, in particular residue E104. σR1.2 has previously been shown to play a role in formation of stable open promoter complexes, in particular it is implicated in allosteric control of -10 promoter element recognition by σR2.3 (Baldwin and Dombroski, 2001, Zenkin et al., 2007). Furthermore, σR1.2 interacts with the 'discriminator' region of the non-template strand of promoter DNA to further stabilise RNAP promoter complexes (Haugen et al., 2006, Haugen et al., 2008b). The dependence of Urd inhibition on σR1.2 integrity suggests the compound mediates inhibition of transcription by interacting with or influencing this particular σ factor sub-region. The preclusion of recognition of the -10 promoter element by Urd suggests the inhibitor may target a crucial allosteric signal from σR1.2 to σR2.3.

Mapping of the Urd resistance mutation E104Q onto the structure of *E. coli* holoenzyme shows the residue is located in the centre of the σR1.2 α-helix (Figure 5.6, B, F). It has been shown previously that holoenzymes with an E104V mutation are unable to form stable open promoter complexes and are highly defective in both abortive and run-off transcription, suggesting this residue is fundamental to the σR1.2 'allosteric signal' (Baldwin and Dombroski, 2001) . Indeed, structural data of the *E. coli* holoenzyme shows this particular residue interacts with residue I290 of the β' coiled-coil domain both within RPo and *apo* form of RNAP (Figure 5.6, F). Concurrently, residue E381 of σR2.1 also makes contacts with the same I290 residue, creating a mechanical linkage between σR1.2 and σR2.1 (Figure 5.6, F). This network of structural contacts is ideally positioned to influence -10 promoter element binding by σR2.3 and it is

reasonable to suggest this region must adopt a strict conformation for successful DNA binding. Interestingly, it has been previously shown that certain mutation of σR1.2 can augment the activity of Fdx at RNAP (Morichaud et al, 2016). However, here, we have shown an effect to the contrary for inhibition by Urd, wherein mutation of σR1.2 decreases activity of the compound at RNAP Nonetheless, although these observations serve to obfuscate interpretation of the data, they serve to reaffirm the role of σR1.2 in inhibition of RPo formation. Indeed, if the Urd resistance mutation E104Q delineates a residue targeted by Urd, it is plausible to propose Urd may effect the structural elements of the σR1.2 allosteric signal and consequently prevent the conformation required for promoter recognition by σR2.3. However, it is possible the compound binds at a site away from these structural elements, and the mutation at residue E104 disrupts a separate allosteric signal linked to σR2.

In summary, Urd is the first transcription inhibitor that blocks the formation of the open promoter complex by targeting σ factor sub-region 1.2, and provides an exciting new compound for future drug development. Additionally, the identification of a novel putative binding site affords a new target for rational drug design. Urd also provides and interesting tool with which to dissect the process of transcription initiation. However, further structural analysis is required to confirm the exact binding site of Urd.

Chapter 6. Kanglemycin A inhibits transcription by a novel steric occlusion mechanism

6.1 Introduction

The natural product antibiotic Kanglemycin A (KglA) was first isolated from the fermentation broth of the actinomycete bacteria *Nocardia mediterranei* var. *kanglensis 1741-64 (Wang et al., 1988)*. The compound was shown to possess activity against certain Gram positive bacteria. However, no mechanism of action had been previously described. The compound is an ansamycin antibiotic structurally related to the rifamycins (Figure 6.1, A). However, KglA possesses structural elements not seen in other naturally occurring or semisynthetic rifamycins. In particular, KglA contrasts from other rifamycins through two large substituents present on the *ansa* bridge; a 2,2-dimethylsuccinic acid chain at C20 and a β-O-3,4-O,O' methylene digitoxose moiety at C27. These variations are particularly unusual as synthetic *ansa*-bridge modifications of bioactive rifamycins frequently lead to inactivation of the compound. Consequently, synthetic modifications have focused primarily on the naphthoquinone moiety.

More recently, our industrial collaborators at Demuris™ identified *Amycolotopsis* DEM30355 as a producer of KglA. In an attempt to identify new transcription targeting antibiotics, the strain extract was screened against the *yvgS Bacillus subtilus* reporter strain that has *lacZ* fused to the promoter of the bacterial helicase *HelD* (Figure 6.1, B). This particular promoter was chosen because it is known to be upregulated in response to sub-inhibitory concentrations of the transcription inhibitor rifampicin (Hutter et al., 2004b). Consequently, in a disk assay, rifamycin type transcription inhibitors produce a blue halo at the frontier of the zone of inhibition upon X-gal infused agar plates. DEM30355 produced a compound (later identified as KglA) that activated this reporter strain. Interestingly, KglA was also active against the Rif resistant *B. subtilus* reporter strain possessing a βH482R (βH526R in *E. coli* numbering) RNAP mutation. Accordingly, DEM30355 was cultivated at 500L scale and KglA extracted and purified by multi-step chromatography. Initial investigative work performed by Dr Lucia Ceccaroni and Dr Hamed Mosaei showed that KglA
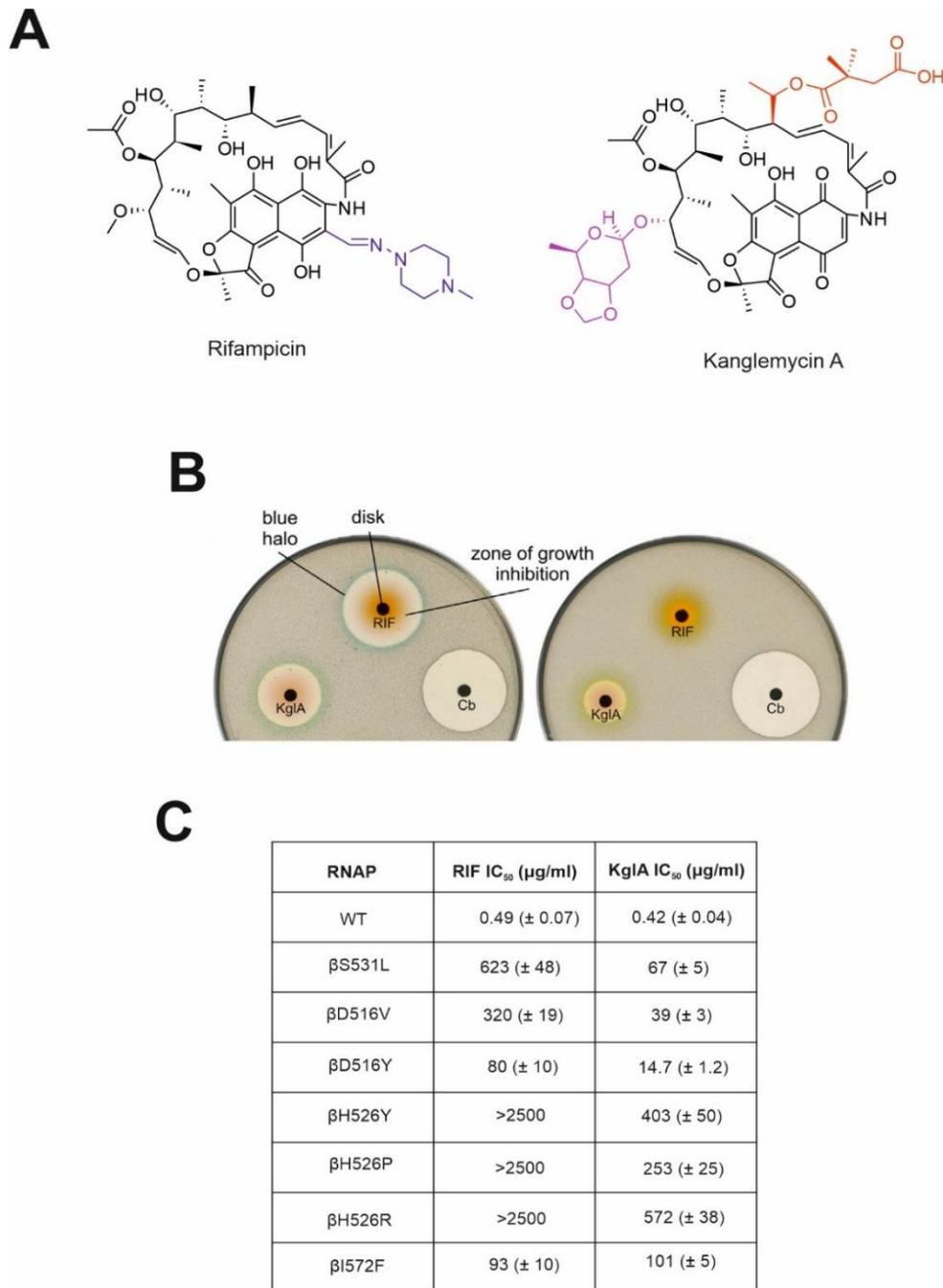
Figure 6.1 Preliminary identification of KglA as an inhibitor of RNAP. Adapted from (Mosaei et al., 2018) (A) Chemical structures of Rifampicin (left), with the synthetic 4-methyl-1-piperazinyl moiety at C3 indicated in purple, and Kanglemycin A (right) with 2,2-dimethylsuccinic acid chain at C20 and a β-O-3,4-O,O' methylene digitoxose moiety at C27, indicated in pink and red, respectively. (B) Disk diffusion assay with Kanglemycin A, Rifampicin, and Carbenicillin (negative control). Paper disks soaked with antibiotic were placed on LB agar plates infused with X-gal and lawn of *yvgS B. subtilis* reporter strain carrying the lacZ gene under the *HelD* promoter (left) and an identical Rif resistant *yvgS B. subtilus* possessing βH482R RNAP mutation (βH526R in *E. coli* numbering) (Mosaei et al., 2018). The *HelD* promoter is induced during partial inhibition of transcription (Hutter et al., 2004b). Note, the blue halo at the frontier of the zone of growth inhibition in the case of transcription inhibitors. (C) Quantification of inhibition of *in vitro* transcription performed by WT *E. coli* RNAP and Rif[R] *E. coli* RNAPs on a linear DNA template

containing the T7A1 promoter. Quantification is derived from average of [^{32}P]-labelled run-off and terminated transcription products. Values were normalised to quantity of [^{32}P] RNA synthesised in the absence of inhibitor. Numbers in brackets are ± SD (Mosaei et al., 2018).

inhibits WT *E. coli* RNAP *in vitro*. Furthermore, the compound retained activity against a suite of Rif-resistant RNAPs including βS531L, βD516V, βD516Y, βH526Y, βH526P, βH526R and βI572F (Figure 6.1, C). These mutants correspond to frequently observed Rif resistant clinical isolates of *M. tuberculosis* (Gill and Garcia, 2011*)*.

These initial results suggest that KglA binds to and inhibits bacterial RNAP. However, the exact mode of action of KglA at RNAP remained unknown. In this study we aimed to elucidate the molecular mechanism of transcription inhibition by KglA.

6.2 Results

6.2.1 Kanglemycin A inhibits transcription at an earlier stage than Rifampicin

Activation of the *helD* promoter by KglA and the compounds structural similarities to rifamycins suggests it may inhibit transcription via a similar mechanism to that of rifampicin. Consequently, we compared the effects of Rif and KglA on transcription by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. In this instance, the dinucleotide primer CpA was used to initiate transcription. Interestingly, KglA inhibited transcription of full length RNA transcripts (both terminated and run off) whilst concurrently causing an accumulation of short tri- and tetra- nucleotide abortive products (Figure 6.2, A). This mechanism is typical of rifamycins (McClure and Cech, 1978, Campbell et al., 2005), as illustrated by transcription inhibition of rifampicin. The principal mode of action of rifamycins is through steric occlusion of the translocating nascent transcript following initial phosphodiester bond synthesis (Campbell et al., 2001).

This results in inhibition of synthesis of full length transcripts but accumulation of short abortive products. Interestingly, however, it was noted that the ratio of tri- (CpApU) to tetra-nucleotide (CpApUpC) abortive products were different for KglA
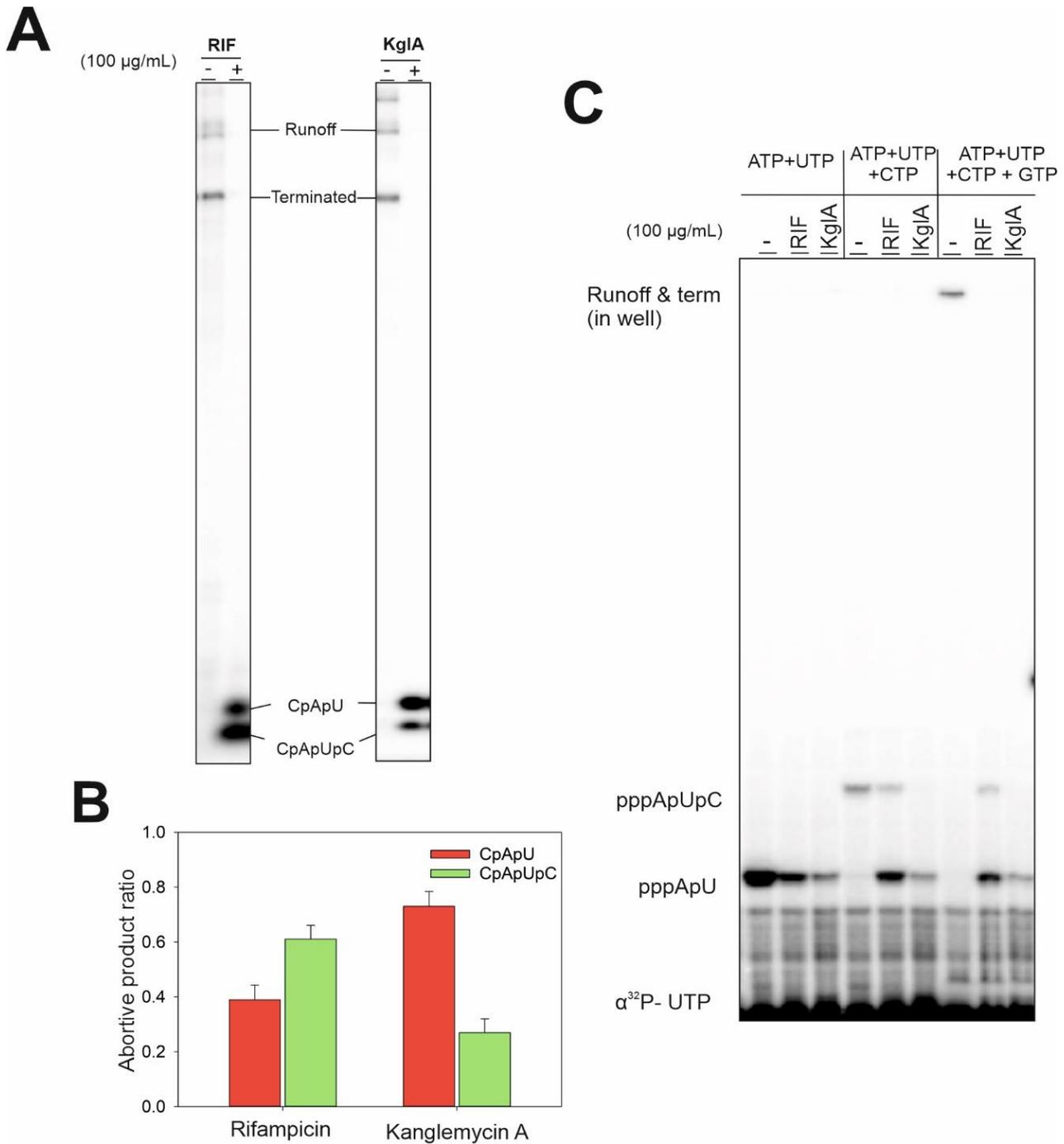
Figure 6.2 Comparable mechanisms of rifampicin (Rif) and kanglemycin A (KglA). (A) *In vitro* transcription by *E. coli* RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of rifampicin (left) or Kanglemycin A (right). [^{32}P]-RNA products (Run-off and terminated) are indicated. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA. [^{32}P]-CpApU and [^{32}P]-CpApUpC are indicated. Note, tetra-nucleotides migrate faster than tri-nucleotides under the electrophoretic conditions used here (Borukhov et al., 1993). (B) Bar chart showing quantification of the ratio of [^{32}P]-CpApU and [^{32}P]-CpApUpC in the presence of Rif or KglA, as shown in (A) - the percentages represent the quantity of the

respective abortive product in respect of the overall quantity of tri- and tetra- nucleotide abortive products synthesised in the respective reaction. (C) *In vitro* transcription by *E. coli* RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of rifampicin or Kanglemycin. Indicated are different ratios of the di-nucleotide (pppApU) and the tri-nucleotide (pppApUpC) abortive products in the presence of Rif and KglA. Transcription was performed in the presence of the nucleotides depicted (in the absence of the CpA primer). Note that, in the 33% gel, runoff and termination products remain in the gel well.

than for Rif (Figure 6.2, B). In the presence of KglA, there is greater inhibition of tetra-nucleotide synthesis compared to Rif. This suggests that KglA acts at a stage preceding that of Rif, perhaps by increased steric hindrance to the translocating nascent transcript.

Interestingly, KglA completely prevents synthesis of triphosphorylated tri-nucleotide and even inhibited the production of triphosphorylated di-nucleotide (~70-fold inhibition), as compared to Rif (~3-fold inhibition) (Figure 6.2, C). This result further indicates that KglA affords additional hindrance to the translocating nascent transcript and may even effect binding of the initiating NTP substrate.

6.2.2 Structural basis of Kanglemycin A – RNAP binding interaction

To reveal the structural basis for RNAP inhibition by KglA, we established a collaboration with Prof Katsuhiko Murakami of the Department of Biochemistry & Molecular Biology at Pennsylvania State University (Penn State). All crystallographic data shown here were gathered by Dr Vadim Molodstov (Penn State) & Prof Katsuhiko Murakami (Penn State).

To analyse binding interactions of the RNAP-KglA complex we determined the X-ray crystal structure of *T. thermophilus* holoenzyme in complex with the *pyrG* promoter template and KglA at 3.0Å resolution (Mosaei et al., 2018) (Figure 6.3, A, B). Electron density maps showed unambiguous density within the Rif binding pocket of the RNAP β subunit. The binding mode of KglA within the binding pocket is highly similar to that of RIF. KglA forms polar contacts with several residues involved in RIF binding; including βQ513, βF514, βS531 and βR540 (*E. coli* numbering) (Figure 6.3, C).

A

B

C

D

Rifampicin
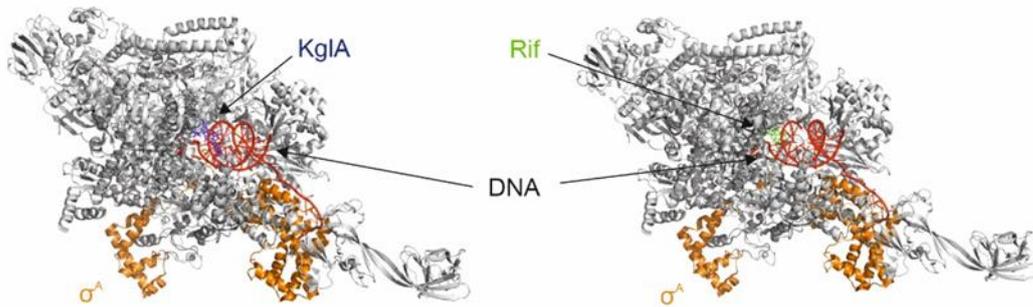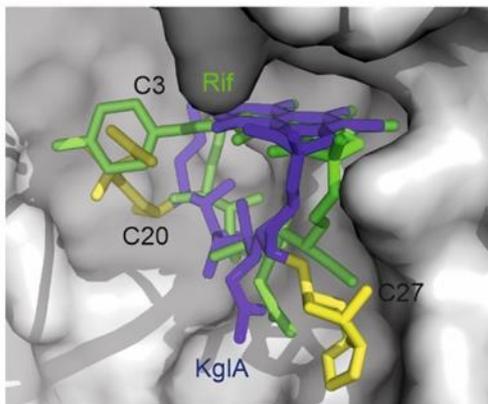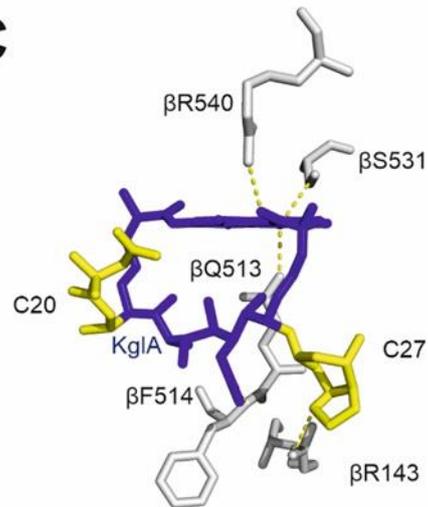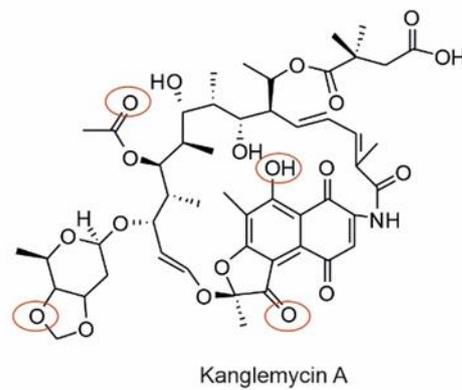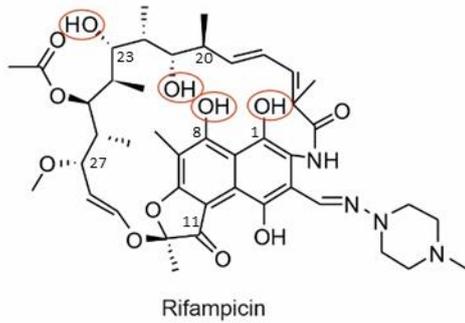
Kanglemycin A

E
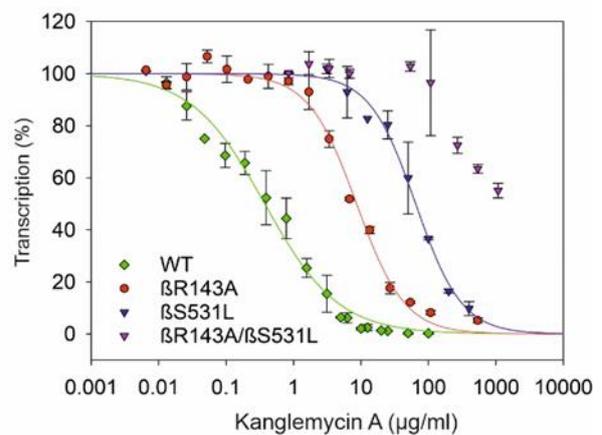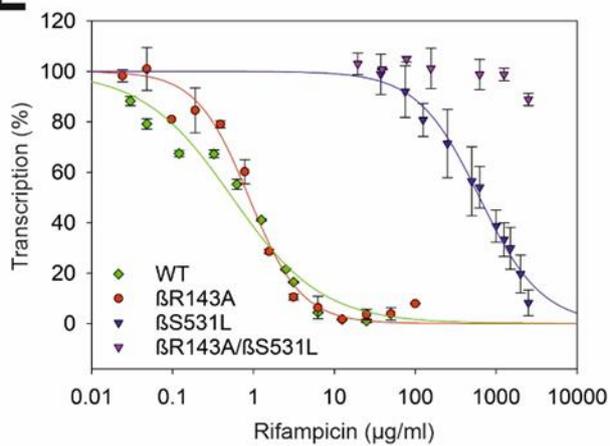
Figure 6.3 Structural basis of kanglemyicn A (KglA) inhibition. (A) Overall view of *T. thermophilus* RNAP with KglA (left, PDB: 6CUU) and Rif (right, PDB: 1YNN) bound within the RIF-binding pocket. RNAP (gray core and orange σ^A) and DNA (red) are depicted as ribbon models, and KglA (blue) and Rif (green) stick model. (B) A close-up view of the RIF-binding pocket of the *T. thermophilus* RNAP containing KglA (in blue, with the side chains in yellow) and Rif (green). RNAP is shown as a semi-transparent surface model (pale gray), KglA is shown as a stick model with the ansa bridge side chains labelled. (C) β subunit residues forming the RIF-binding pocket shown as stick models. The hydrogen bonds between KglA and β amino acid residues are depicted by yellow dotted lines. (D) Critical oxygen functionalities present in rifampicin (left) and kanglemycin (right) involved in binding to RNAP are circled in red. Relevant carbon atom numbering is indicated in the left panel (E) Quantification of inhibition of *in vitro* transcription by rifampicin (left panel) kanglemycin (right panel). Transcription was performed by WT *E. coli* RNAP and mutant RNAPs on a linear DNA template containing the T7A1 promoter. Reactions were initiated with 5'- non-phosphorylated dinucleotide primer CpA. Quantification is derived from average of [^{32}P]-labelled run-off and terminated transcription products. Values were normalised to quantity of [^{32}P] RNA synthesised in the absence of inhibitor. Error bars are ± SD from at least 3 independent experiments.

However, interestingly, KglA conformation within the Rif pocket is subtly different to that of Rif. The *ansa*-bridge adopts a slightly tilted conformation away from the molecular surface of the β subunit preventing the polar interaction of C23 with βF514 seen in RNAP-Rif complexes, an interaction thought essential to Rif binding. However, the β-O-3,4-O,O' methylene digitoxose moiety at C27 creates a new hydrogen bond with the side chain of βR143; this interaction does not occur with Rif (Figure 6.3, C, D). Furthermore, the acid side chain establishes further non-polar interactions with surrounding residues dramatically increasing the overall binding surface of KglA compared to Rif.

To validate this novel interaction by the novel C27 β-O-3,4-O,O' methylene digitoxose moiety with βR143, we created *E. coli* RNAP with βR143A mutation. The mutation had no effect on sensitivity to Rif. However, the mutant RNAP was over 10-fold less sensitive to KglA, reaffirming the importance of this residue in KglA binding. Nevertheless, the IC50 of KglA for the mutant RNAP is still lower than for any Rif resistant mutant RNAP, suggesting other interactions within the binding pocket are still important. Indeed, only the introduction of a βR143A concurrently with the common clinical Rif resistant mutation βS531L induced resistance to KglA (Figure 6.3, E). (Note, *in vitro* transcription data for Rif shown in Figure 6.3 E, left panel, was gathered with support from Dr Hamed Mosaei)

6.2.3 Structural basis of Kanglemycin A mode of action

Rifamycin inhibitors act through steric hindrance of translocating nascent transcript *following* the synthesis of the first or second phosphodiester bond formation (McClure and Cech, 1978). Consequently, this leads to a reduction in full length transcripts and increase in short abortive products. However, during *de novo* transcription, KglA inhibits transcription of both full length transcripts and abortive products, indicating that KglA acts at an earlier stage than Rif.

To further investigate the effects of KglA on substrate binding and phosphodiester bond formation, the structure of a *T. thermophilus* RNAP *de novo* initiation complex (4Q4Z) was overlaid onto the *T. thermophilus* RNAP-KglA and RNAP-Rif complex structures. Modelling of Rif shows no substantial steric clash with initiating substrate nor DNA. Additionally, the piperazine N4 is positively charged and thus will not disfavour NTP binding in the i site through electrostatic repulsion with the negatively charged oxygens of the γ-phosphate (Figure 6.4, A). This is consistent with previous data suggesting Rif has a minor effect on K_m of initiating substrate (McClure and Cech, 1978). Furthermore, there is sufficient flexibility in the Rif C3 sidechain and also sufficient space to accommodate extension of the nascent transcript to 3-mer before a significant clash with the γ-phosphate of the initiating nucleotide.

Structural modelling of KglA, however, within *T. thermophilus* RNAP *de novo* initiation complex indicates that the orientation of the unique C20 2,2-dimethylsuccinic acid moiety would generate a moderate steric clash with the γ-phosphate of the initiating NTP (Figure 6.4, B). This moderate clash explains how KglA can inhibit dinucleotide synthesis. Additionally, the modelled structure shows the negatively charged carboxylic acid group of the 2,2-dimethylsuccinic acid would further disfavour initiating NTP binding through electrostatic repulsion of the negatively charged oxygens of the 5'-triphosphate group (Figure 6.4, B).

To test this hypothesis, we assessed KglA inhibition of transcription initiated with nucleotide monophosphate (NMP). Structural data suggest there would be no *steric* clash between the C20 2,2-dimethylsuccinic acid moiety and the α-phosphate of an initiating NMP. However, *electrostatic* repulsion between 2,2-dimethylsuccinic acid side chain and α- phosphate may disfavour NMP binding

and transcript translocation and consequently inhibit di- and tri-nucleotide synthesis. As can be seen from Figure 6.4, C, as compared to Rif, KglA moderately inhibits synthesis of pApU suggesting this electrostatic repulsion plays an influential role in KglA inhibition. Furthermore, synthesis of pApUpC (which is not inhibited by Rif) is strongly inhibited by KglA. This data reaffirms both steric and electrostatic effects of KglA hinder both formation of the first phosphodiester bond and translocation of the 5'-phosphorylated dinucleotide.
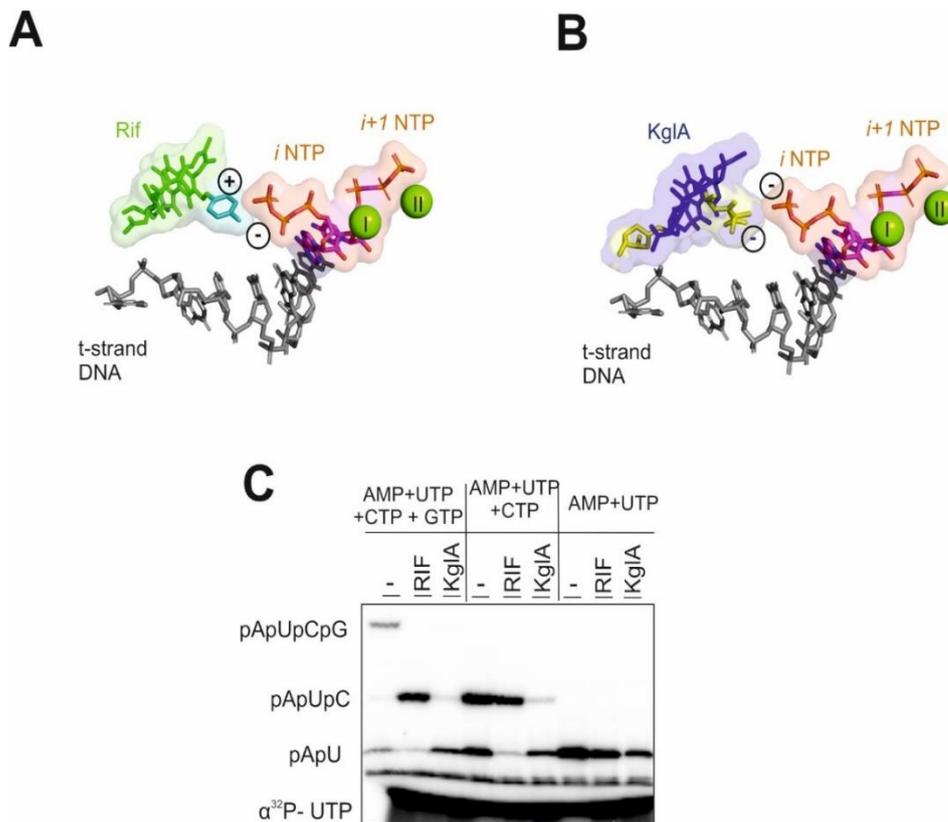
Figure 6.4 Structural basis for mechanism of action of kanglemycin A. View of the RNAP active site from *T. thermophilus de novo* initiation complex (PDB: 4Q4Z) with bound Rif (green) (PDB: 1YNN) superimposed. Depicted is the t-strand DNA from +1 to −7 (light gray), the initiating NTP substrates (i site NTP, ATP; *i*+1 NTP, CMPcPP; (orange/purple)) ,and two Mg^{2+} ions , Mg-I and Mg-II (green spheres; Mg-I is the bound in the active site, Mg-II is bound to the *i*+1 NTP). Rif and the NTPs are depicted as transparent surface models. Positive charge (+) of the Rif piperazine moiety (blue), and negative charge (-) of the *i* NTP γ-phosphate are indicated. (B) Identical to (A) but depicting Kanglemycin A (coloured as in Fig. 5.3). Note the proximity of electrostatic charge in each case. (C)) *In vitro* transcription by *E. coli* RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of rifampicin or Kanglemycin. Indicated are different ratios of the mono-phosphorylated di-nucleotide (pApU) and mono-phosphorylated tri-nucleotide (pApUpC) abortive products in the presence of RIF and KglA. Transcription was performed in the presence of the nucleotides depicted (in the absence of the CpA and ATP primer).

A

Organism/isolate	Strain Reference	RIF- Resistance	RIF MIC (µg/ml)	KglA IC$_{50}$ (µg/ml)
M.tuberculosis	H37Rv	susceptible	0.005 (± 0.001)	0.33 (± 0.09)
M.tuberculosis	1192/015	susceptible	0.004 (± 0.001)	0.21 (± 0.07)
Multi-drug-resistant *M.tuberculosis*	08/00483E	resistant	>4.096	0.85 (± 0.02)

B

RRDR-N

M. tuberculosis	167	RVVSQ
M. smegmatis	164	RVVSQ
E. coli	143	RVISQ
T. thermophillus	134	RVISQ

RRDR-I

M. tuberculosis	423	EFRGTSQLSQFMDQNPLSGLTHKRRLSALGPGGISR
M. smegmatis	420	EFRGTSQLSQFMDQNPLSGLTHKRRLSALGPGGISR
E. coli	504	EFRGSSQLSQFMDQNPLSEITHKRRISALGPGGITR
T. thermophillus	384	EFRSRSQLSQFKDENPLSSLRHKRRISALGPGGITR

RRDR-II

M. tuberculosis	481	ETPEGPNIGLEGSL
M. smegmatis	478	ETPEGPNIGLEGSL
E. coli	562	ETPEGPNIGLENSL
T. thermophillus	442	ETPEGANIGLETSL

RRDR-III

M. tuberculosis	604	NMQRQAVP
M. smegmatis	601	NMQRQAVP
E. coli	684	NMQRQAVP
T. thermophillus	563	NMQTQAVP

C

Rifampicin

Kanglemycin A

(µg/mL) - 12.5 6.25 3.125 1.563 0.781 0.391 0.195 0.098 0.049 0.024 0.012 0.006 0.003 0.0015 0.0007

Runoff
Terminated
CpApU
CpApUpC
α^{32}P- UTP

D

ATP+UTP | ATP+UTP +CTP

(100 µg/mL) - RIF KglA - RIF KglA

pppApUpC
contamination from α^{32}P- UTP
pppApU
α^{32}P- UTP

E

M. smegmatis RNA polymerase

Transcription (%)

Inhibitor (µg/ml)

Inhibitor	IC$_{50}$ (µg/ml)
KglA	0.043 (±0.004)
RIF	0.011 (±0.001)

Figure 6.5 Kanglemycin inhibits *Mycobacterial* RNAP and retain activity against Rif[R] *M. tuberculosis*. (A) Table indicating the MIC values of Rif-susceptible *M. tuberculosis* strains H37Rv and 1192/015, and RIF-resistant *M. tuberculosis* 08/00483E by KglA and RIF. Data presented are the mean of four independent experiments ± SD (Mosaei et al., 2018). (B) Sequence alignment of the Rif resistance determining regions (RRDRs) of the β subunit of RNAPs from *M. tuberculosis, M. smegmatis*, *E. coli* and *T. thermophilus*. (C) *in vitro* transcription performed by WT *M. smegmatis* RNAP on a linear DNA template containing the T7A1 promoter in the presence and absence of rifampicin and kanglemycin A. [^{32}P]-RNA products (Run-off and terminated) are indicated. Reactions were initiated with the 5'-non-phosphorylated dinucleotide primer CpA. (D) *In vitro* transcription by *M. smegmatis* RNAP on a linear DNA template containing the T7A1 promoter, performed in the absence or presence of rifampicin or Kanglemycin. Indicated are different ratios of the di-nucleotide (pppApU) and the tri-nucleotide (pppApUpC) abortive products in the presence of Rif and KglA. Transcription was performed in the presence of the nucleotides depicted (in the absence of the CpA primer) (E) Quantification of inhibition of *in vitro* transcription by rifampicin (left panel) kanglemycin (right panel). Transcription was performed by WT *M. smegmatis RNAP* on a linear DNA template containing the T7A1 promoter. Reactions were initiated with 5'- non-phosphorylated dinucleotide primer CpA. Quantification is derived from average of [^{32}P]-labelled run-off and terminated transcription products. Values were normalised to quantity of [^{32}P] RNA synthesised in the absence of inhibitor. Error bars are ± SD.

6.2.4 Kanglemycin A is active against Mycobacterial RNAP and MDR *M. tuberculosis*

Rifampicin has been used as a frontline treatment for over half a century against infections caused by *M. tuberculosis.* However*,* Rif resistant strains are becoming increasingly prevalent and problematic to treat (Zumla et al., 2015). Our data thus far suggests KglA may be effective in treating Rif resistant bacteria, in particular MDR-*M. tuberculosis*. To investigate the effectiveness of KglA against MDR- *M. tuberculosis* we entered into a collaboration with Dr Joanna Bacon of the TB research group at Public Health England. MIC determination data shown here were gathered by Dr Joanna Bacon.

To assess activity against Rif resistant *M. tuberculosis* we determined MIC values for Rif and KglA against a clinical MDR-*M. tuberculosis* isolate, Beijing 08/00483E. This particular isolate is fully resistant to Rif and is also resistant to all other first-line antitubercular drugs (ethambutol, isoniazid and pyrazinamide). Beijing 08/00483E carries the most frequently observed amino acid mutation in Rif resistant *M. tuberculosis* (βS450L, *M. tuberculosis* numbering; S531L, *E. coli* numbering) (Jamieson et al., 2014, Zhang et al., 2014b). In addition to the MDR-isolate, we also assessed activity against two Rif susceptible strains, H37Rv and 1192/015. Expectedly, the MDR-resistant isolate Beijing 08/00483E is several

orders of magnitude less sensitive to Rif than the Rif-sensitive strains (Figure 6.5, A). On the contrary, KglA was active against all 3 strains in a dose-dependent manner, demonstrating KglA is active against MDR-*M. tuberculosis.* Despite this, KglA was less active than Rif against Rif-sensitive strains. This is apparently contradictory to *in vitro* data suggesting KglA and Rif are equally active against *E. coli* RNAP (Figure 6.1, C). It is plausible KglA and Rif act differently at *M. tuberculosis* RNAP. Subsequently, we assessed transcription inhibition by Rif and KglA at *M. smegmatis* RNAP, a commonly used experimental model for *M. tuberculosis* RNAP. The amino acid composition of the Rif-binding pocket is identical in both Mycobacterial RNAPs (Figure 6.5, B). The mode of transcription inhibition by KglA and Rif at *M. smegmatis* appeared very similar to inhibition of *E. coli* RNAP. At high concentrations both Rif and KglA inhibited full length transcripts with a concurrent accumulation of short abortive RNAs (Figure 6.5, C). Again, as with *E.coli* RNAP, synthesis of CpApUpC was observed in the presence of Rif but inhibited in the presence of KglA (transcription was initiated with CpA) (Figure 6.5, C). This is indicative KglA, when compared to Rif, retains its additional hindrance to translocation of 5' non-phosphorylated nascent transcript at Mycobacterial RNAP. Indeed, the mode of inhibition of *de novo* transcription by KglA was also identical to that seen at *E.coli* RNAP (Figure 6.5, D). Synthesis of pppApU was strongly inhibited by KglA, yet Rif allowed synthesis of dinucleotide triphosphate and even permitted some synthesis of pppApUpC. The IC_{50} of KglA at *M. smegmatis* RNAP indicated it was around four-fold less active than Rif (Figure 6.5, E). It is worthy of mention, the increased sensitivity of Mycobacterial RNAP to rifamycins is well documented in prior literature (Zenkin et al., 2005b). Despite the slightly lower activity of KglA at mycobacterial RNAP when compared to Rif, the respective activities cannot explain the disparity in activity between Rif and KglA at Rif-sensitive *M. tuberculosis.* The most plausible explanation for this difference is due to decreased ability of KglA to cross the cell envelope.

Discussion

The growing prevalence of Rif resistant infections demands compounds with novel modes of action and novel binding interactions. This study has shown KglA forms previously undescribed binding contacts within the RNAP Rif binding

pocket and inhibits transcription through a novel molecular mechanism. Structural analysis of KglA bound within *T. thermophilus* open promoter complexes has shown the additional sugar (β-O-3,4-O,O' methylene digitoxose) and acid (2,2-dimethylsuccinic acid) moieties at C27 and C20, respectively, establish new contacts within the Rif binding pocket and influence compound conformation within the Rif binding pocket. The slight rotation away from the binding pocket may allow the compound to attenuate the effect of certain resistance mutations (Mosaei et al., 2018). Consequently, the altered conformation and novel binding interactions within the Rif pocket may explain the preliminary data indicating activity of KglA against a suite of Rif-resistant bacterial RNAPs. Indeed, our data here show only a double mutation within the Rif binding pocket leads to resistance *in vitro*.

The mode of action of KglA is different to that of Rif. During *de novo* transcription, KglA inhibits the formation of the first phosphodiester bond, whereas Rif inhibits the formation of the second phosphodiester bond. With our collaborators, the combined biochemical and structural approach we have taken here has allowed us to discern the basis of this mechanistic difference. The C20 2,2-dimethylsuccinic acid moiety appears to afford an additional steric and electrostatic obstacle to both initiating nucleotide binding and to translocation of the nascent transcript. Interestingly, inhibition of first phosphodiester bond synthesis has been described previously for the semisynthetic rifamycin, rifabutin (Artsimovitch et al., 2005). However, in this instance, 5' non-phosphorylated dinucleotide ApU (corresponding to +1,+2 positions) was used to initiate transcription from the T7A1 promoter, unlike our use here of CpA (corresponding to -1,+1 positions). Consequently, our results are not comparable. Nevertheless, our data show KglA failed to inhibit synthesis of 5' non-phosphorylated trinucleotide, despite the earlier projected clash with 5' cytidine when transcripts are initiated with CpA, suggesting they likely act through different mechanisms.

KglA inhibition of first phosphodiester bond formation is likely to offer a further ancillary mechanism of inhibition, when compared to Rif. RNAP can continually synthesis 2-3 nucleotide abortive products (termed RNA priming) when bound to Rif (McClure and Cech, 1978, Campbell et al., 2001). However, Rif binding is precluded when RNAs become longer than 3-mer. Thus, the likelihood of RNAP

synthesising a longer transcript, which would then prevent Rif binding if Rif were to transiently dissociate, is quite probable. However, inhibition of first phosphodiester bond formation by KglA would serve to prevent this RNA priming. This effect may be of particular relevance in the context of Rif-resistant RNA polymerases in which binding affinity will be significantly reduced, and the transient dissociation of antibiotic becomes more probable (Molodtsov et al., 2017b, Campbell et al., 2001).

Semi synthesis of the rifamycin class has generally centred on variations of the C3/C4 positions of the naphthoquinone moiety (Bacchi et al., 1998, Sensi, 1983). Synthetic modifications of the *ansa*-bridge have generally rendered compounds inactive or significantly less potent. The unique substituents of KglA described here reveal the significant potential of *ansa*-bridge modifications, particularly the novel binding interactions that can be utilised by substitution at the C27 position. Furthermore, flexible acidic substitutions at the C20 position can facilitate novel interactions with the initiating nucleotide. It appears, however, that these native modifications come at the cost of reduced penetration of the cell envelope. This is most likely a consequence of the polar carboxyl groups present on the C20 2,2-dimethylsuccinic acid sidechain. Indeed, despite a highly similar potency and mode of action at mycobacterial RNAP, KglA has a significantly reduced potency against Rif-susceptible strains of *M. tuberculosis* when compared to Rif. Although, it should be noted, KglA still possesses MIC values below those of frontline antitubercular drugs isoniazid and pyrazinamide. Importantly, however, KglA remained active against the MDR-*M. tuberculosis* isolate, Beijing 08/00483E, which is completely resistant to Rif. The struggle to find compounds active against MDR-*M. tuberculosis,* plus the beneficial mechanistic characteristics of the inhibitor we have described here, demand KglA be optimised for clinical usage.

In summary, KglA is a naturally evolved solution to the question of Rif-resistant bacteria, in particular Rif-resistant *M. tuberculosis.* Our study has illustrated there is still great scope to effectively explore the chemical space surrounding the rifamycin core. In particular, novel binding interactions within the Rif pocket can be utilised by substitution at the C27 position. Furthermore, substitution at C20 can exploit novel inhibitory mechanisms not typically seen in the rifamycin class.

Namely, additional steric and electrostatic hindrance to translocating nascent RNA and binding of initiating nucleotide. Together, KgIA offers an exciting new lead towards effective therapeutic agents.

Chapter 7. Rifampin ADP-ribosyl transferases from *M. smegmatis* and *M. abscessus* have differing substrate specificities.

7.1 Introduction

Resistance to the front line TB drug Rifampicin (Rif) is typically conferred through point mutations within the *rpoB* gene encoding the Rif binding β-subunit of RNAP (Campbell et al., 2001). However, Rif is also subject to enzymatic inactivation by several recently described bacterial enzymes (Spanogiannopoulos et al., 2012, Baysarowich et al., 2008, Liu et al., 2016, Stogios et al., 2016). ADP-ribosylation of Rif by *M. smegmatis* Rifampicin ADP-ribosyl transferase (Arr$_{Ms}$) is believed to preclude Rif binding to the RNAP β subunit. Arr$_{Ms}$ utilises an NAD+ cosubstrate to catalyse the ADP-ribosylation of Rif at the C23 hydroxyl group, with concurrent expulsion of the nicotinamide moiety from NAD+ (Morisaki et al., 2000) (Figure 7.1, A). Rifamycin SV, and newer semi-synthetic Rif derivatives rifaxamin and rifabutin, are also substrates for Arr$_{Ms}$, suggesting the enzyme has a broad substrate specificity (Baysarowich et al., 2008). However, it has recently been shown C25 carbamate derivatives of rifampicin exhibit greatly improved antimicrobial activity against *M. smegmatis* (Combrink et al., 2007). *In vitro* experiments with purified Arr$_{Ms}$ indicated C25 carbamate rifamycins are resistant to inactivation by Arr (Combrink et al., 2007). The recently solved crystal structure of Arr$_{Ms}$ in complex with Rif provides a rational explanation for this evasion of Arr inactivation; the supposed position of the large C25 carbamate group would afford a pronounced clash within the Arr$_{Ms}$ Rif binding pocket (Baysarowich et al., 2008). Seemingly, C25 carbamates do not bind to the enzyme and evade inactivation by ADP-ribosylation. It was recently discovered the genome of the pathogen *M. abscessus* may encode a Rifampicin ADP-ribosyl transferase (Arr$_{Mab}$) conferring innate high-level Rif resistance. Deletion of Arr$_{Mab}$ increases sensitivity of *M. abscessus* to Rif by over 500-fold. Interestingly, C25 modified rifamycins (Figure 7.1, C) showed increased activity against WT *M. abscessus,* but also against an *M. abscessus* mutant lacking Arr$_{Mab}$, suggesting increased activity of C25 derivatives may not be due to resistance to Arr$_{Mab}$ inactivation in *M.*

abscessus (Rominski et al., 2016). This indicates the putative Arr_{Mab} may have a different substrate specificity to Arr_{Ms}.
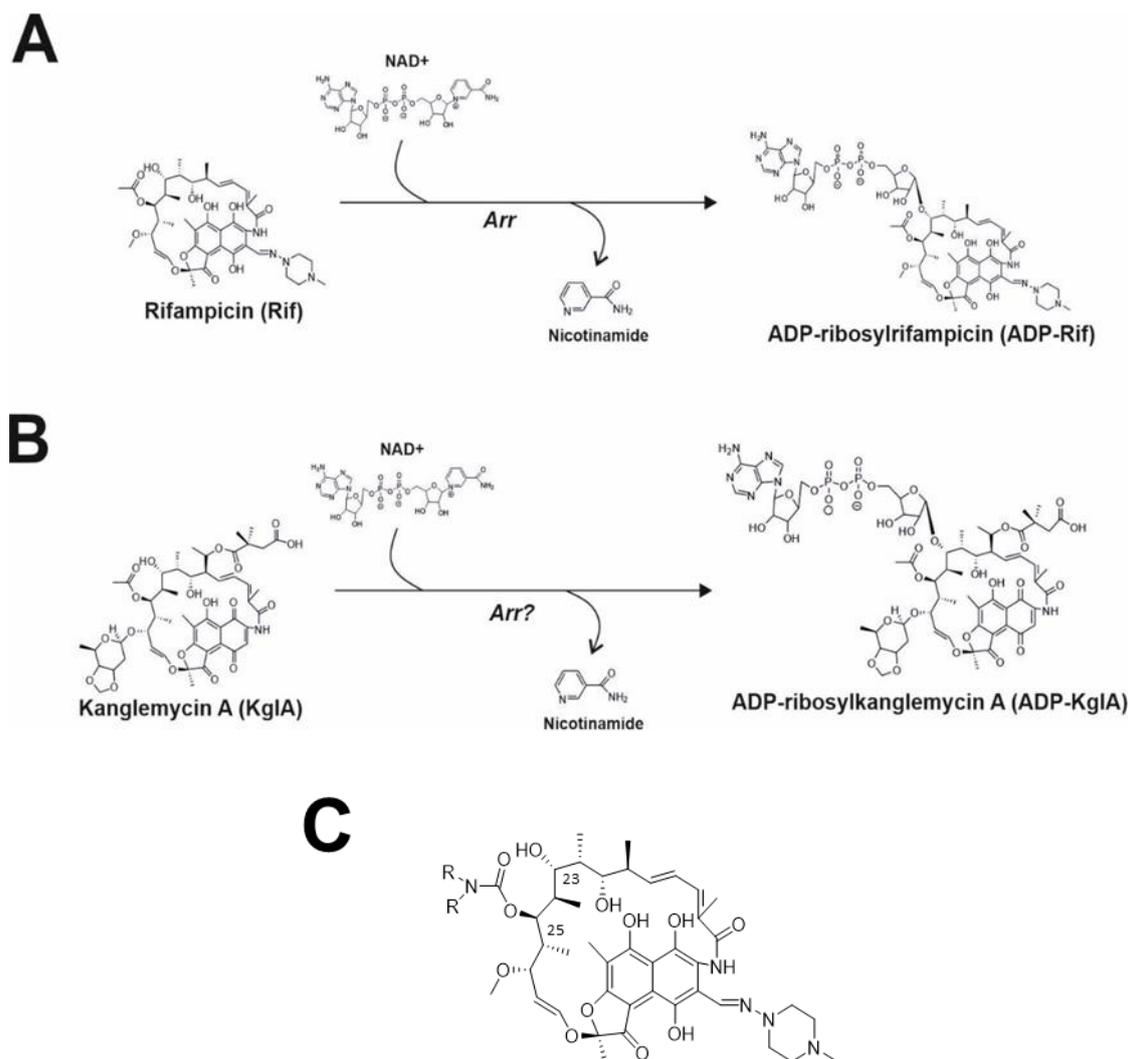
Figure 7.1 Structural scheme of ADP-ribosylation by rifampicin ADP ribosyltransferases (Arr) (A) ADP-ribosylation of rifampicin by Mycobacterial Arr (Baysarowich et al., 2008) (B) Anticipated ADP-ribosylation of kanglemycin A by Arr. (C) The chemical structure of C25 carbamate rifampicin derivatives.

The ability of C25 carbamates to avoid inactivation by Arr_{Ms} suggests other modifications to the *ansa-* chain may preclude binding to Arr. The natural product KglA is an ansamycin antibiotic that inhibits transcription by binding within the Rif binding pocket on the β-subunit of RNAP. The compound has distinctive substituents present on the *ansa* bridge; a 2,2-dimethylsuccinic acid chain at C20 and a β-O-3,4-O,O' methylene digitoxose moiety at C27. When bound to RNAP, these substituents afford additional binding contacts in the Rif pocket and

103

produce a distinctivive binding conformation (Mosaei et al., 2018). Consequently, KglA inhibits RNAP by a unique steric occlusion mechanism and retains activity at Rif resistant RNAPs. Whether these unique moieties effect KglA binding to Arr is yet to be assessed. We speculated these large bulky substituents may prevent KglA binding to Arr in a manner analogous to C25 carbamate Rif derivatives (Figure 7.1, B). Consequently, in this study, we aimed to discern if *M. abscessus Arr* encodes a functional Rifampicin ADP-ribosyl transferase, and aimed to determine if KglA is a substrate for purified Arr_{Ms} and Arr_{Mab}. We also assessed the activity of ADP-ribosylated rifampicin and ADP-ribosylated KglA in an *in vitro* transcription assay.

7.2 Results

7.2.1 Characterisation of *Mycobacterium smegmatis* Arr (Arr_{Mab})

Firstly, in order to characterise Arr_{Ms} *in vitro*, we cloned the *Arr* gene from *M. smegmatis* into a pET28 expression vector, and expressed and purified the protein from *E.coli*. To assess activity of Arr_{Ms}, we utilised a tandem LC-MS based assay in which antibiotic substrate and NAD+ cosubstrate were incubated in the absence and presence of Arr_{Ms}, and then the reaction products separated by HPLC and resolved by tandem mass spectrometry. When Rif is incubated with NAD+ in the absence of Arr_{Ms}, Rif, and its oxidised form rifampicin quinone (Rifq), are resolved with no identifiable ADP-ribosylated product (Figure 7.2, A, B). However, when Rif is incubated in the presence of NAD+ and Arr_{Ms}, both Rif and Rif quinone are ADP-ribosylated, confirming the activity of our purified Arr_{Ms} (Figure 7.2, top panel). We then assessed if KglA was subject to ADP-ribosylation by Arr_{Ms}. As seen previously with Rif, incubation of KglA with NAD+ in the absence of Arr_{Ms} afforded no ADP-ribosylation of the antibiotic. Interestingly, additional incubation with Arr_{Ms} also failed to ADP ribosylate KglA, suggesting that KglA is not a substrate of Arr_{Ms} (Figure 7.2, bottom panel).

To clarify the *in vivo* efficiency of antibiotic inactivation, or lack thereof, of Arr_{Ms}, we performed *in vitro* inactivation reactions with a tandem disk assay. Increasing concentrations of Arr_{Ms} were coincubated with NAD+ and antibiotic (1mg/ml), and then spotted onto paper disks which were placed onto a lawn of *S. aureus* RM4220. Apparent decreases in zones of inhibition were interpreted to signify

functional inactivation of the respective antibiotic. As shown in figure Figure 7.2, E, the control antibiotic carbenciliin is not inactivated by Arr$_{Ms}$. However, as indicated by our LC-MS based assay, rifampicin is inactivated by *Arr$_{Ms}$,* resulting in decreased zones of inhibition when Arr$_{Ms}$ concentration is increased. Additionally, consistent with our previous data, KglA is resistant to inactivation by Arr$_{Ms}$, as indicated by no decrease in zone of inhibition even at very high concentrations of Arr$_{Ms}$. Combined with previous data, this result reaffirms KglA is not a substrate for
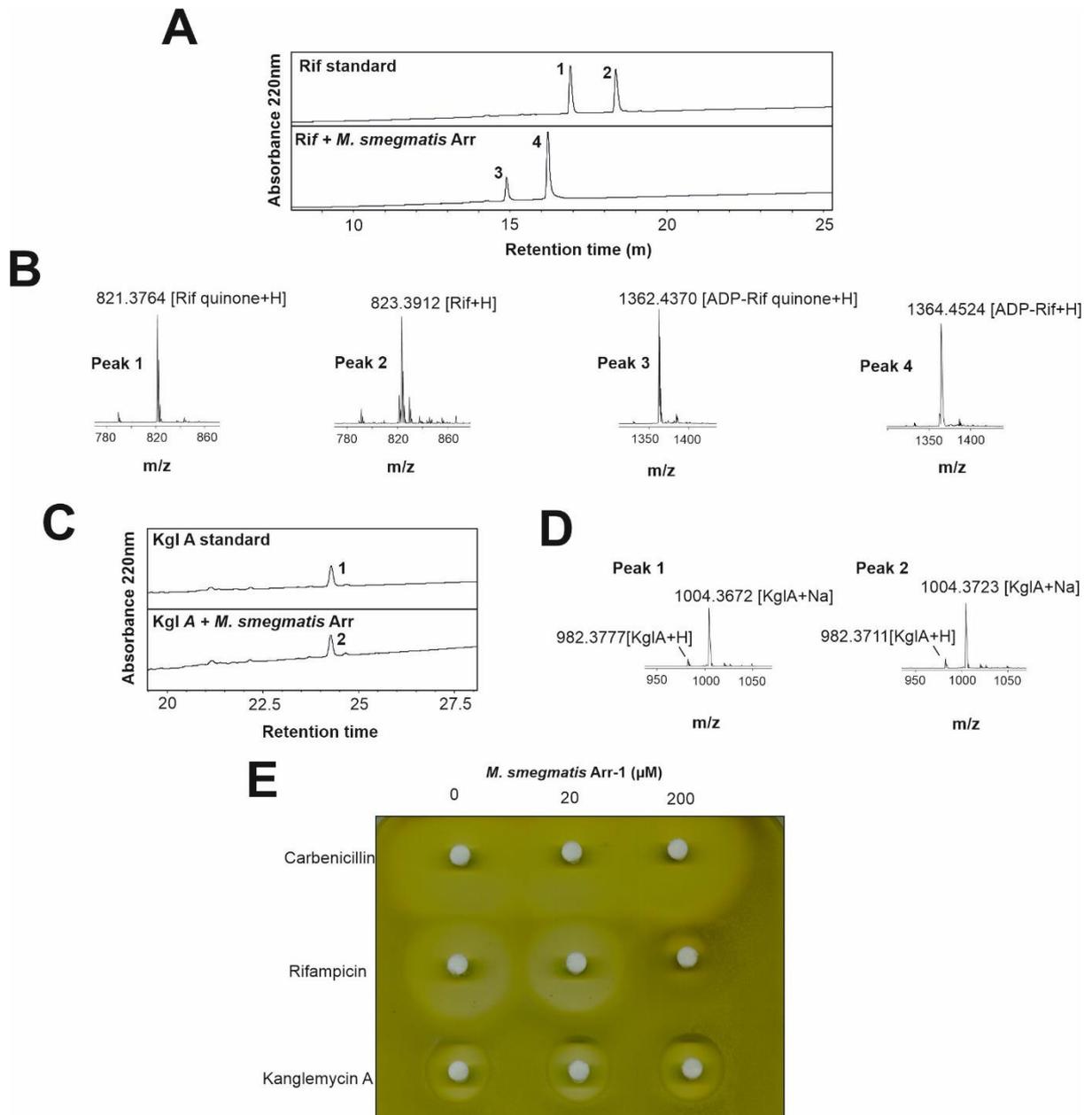
Figure 7.2 *M. smegmatis* Arr (Arr$_{ms}$) fails to inactivate kanglemycin A (KglA) by ADP-ribosylation. Representative reverse phase HPLC traces of *in vitro* inactivation of rifampicin (Rif) by Arr$_{ms}$. (Upper panel) Trace represents the Rif standard in which Rif is incubated with NAD+ in the absence of Arr$_{Ms}$. (Lower panel) Trace represents Rif incubated with NAD+ for 1h in the presence of Arr$_{Ms}$. Absorbance is at 220nm. (B) Positive ion mass spectrum corresponding to peaks indicated '1', '2', '3' and '4' identified in (A). Ion adducts are as indicated. (C) Representative reverse phase HPLC traces of *in vitro* incubation of KglA (Rif) with Arr$_{ms}$. (Upper panel) Trace represents the KglA standard in which Rif is incubated with NAD+ in the absence of Arr$_{Ms}$. (Lower panel) Trace represents KglA incubated with NAD+ for 1h in the presence of Arr$_{Ms}$. Absorbance is at 256nm. (D) Positive ion mass spectrum corresponding to peaks indicated '1' and '2' identified in (C). Probable Ion adducts are as indicated. (E) Disk diffusion assay of aliquots of *in vitro* reactions (corresponding to (A) and (C)) with carbenicillin, rifampicin, or kanglemycin A incubated in the presence of increasing concentrations of Arr$_{Ms}$. Paper disks were soaked with aliquots from *in vitro* reactions and were placed on LB agar plates infused with lawn of *S. aureus* RM4220 (Combrink et al., 2007).

Arr$_{Ms}$. Note the smaller zones of inhibition seen with KglA, when compared to Rif, is thought to be due to poorer penetration of the compound through the cell envelope.

Furthermore, our collaborators in the lab of Prof. Peter Sander at the University of Zurich have gathered MIC data for Rif and KglA with both WT *M. smegmatis* and *ΔArr$_{Ms}$ M. smegmatis* (unpublished)*.* Their results show *ΔArr$_{Ms}$ M. smegmatis* is considerably more sensitive to Rif than the WT strain. However, there was essentially no difference in sensitivity of both WT and *ΔArr$_{Ms}$* strains to KglA, further corroborating KglA is not a substrate for Arr$_{Ms}$.

7.2.2 Characterisation of *Mycobacterium abscessus* Arr (Arr$_{Mab}$)

Prior genetic experiments have indicated the substrate specificity of a putative Arr$_{Mab}$ may differ from that of Arr$_{Ms}$. Consequently, we aimed to determine the function and activity of Arr$_{Mab}$ *in vitro*. We cloned the *Arr* gene from *M. abscessus* into a pET28 expression vector, and expressed and purified the protein from *E. coli*. Once more, we utilised our tandem LC-MS based assay to determine the activity of Arr$_{Mab}$. As previously, coincubation of Rif and Rif quinone with NAD+ in the absence of Arr$_{Mab}$ resulted in no ADP-ribosylated products (Figure 7.3, top panel). However, upon coincubation of NAD+ and Rif with Arr$_{Mab}$ we see a shift in retention time distinctive of ADP-ribosylation of Rif and Rif quinone. ADP-ribosylation of Rif was confirmed through identification of masses corresponding to inactivated products (Figure 7.3, top panel). This result confirms that the

apparent *Arr* gene present in *M. abscessus* encodes a functional Rifampicin ADP-ribosyl transferase. Next, we attempted to determine if KglA is a substrate for Arr$_{Mab}$. As previously, coincubation of KglA and NAD+ in the absence Arr$_{Mab}$ fails to facilitate ADP-ribosylation of the antibiotic (Figure 7.3, bottom panel). However, when KglA is incubated with Arr$_{Mab}$ and the cosubstrate NAD+ there is a characteristic shift in retention time of the product, analogous to that seen with rifampicin. The peak corresponds to a compound with the exact mass of 1523.40 Daltons; the predicted mass of ADP-ribosyl KglA (Figure 7.3, bottom panel). This confirms Arr$_{Mab}$ can inactivate KglA, and possesses a broader substrate specificity than its homologue Arr$_{Ms}$.

To ascertain the efficacy of antibiotic inactivation by Arr$_{Mab}$ within an *in vivo* context, as for Arr$_{Ms}$, we performed *in vitro* inactivation reactions alongside tandem disk assay. As previously, increasing concentrations of Arr$_{Mab}$ were coincubated with NAD+ and antibiotic, and then spotted onto paper disks which were placed onto lawns of *S. aureus* RN4220. Again, apparent decreases in zones of inhibition were interpreted to indicate functional inactivation of the respective antibiotic. As suggested by our LC-MS assay, Arr$_{Mab}$ is a functional Rifampicin ADP-ribosyl transferase. Rif activity is almost completely eliminated by both low- and high- range concentrations of Arr$_{Mab}$ tested here (Figure 7.3, E). Fascinatingly, Arr$_{Mab}$ is indeed capable of inactivating KglA; there was marked reduction in the zone of inhibition when KglA is incubated with Arr$_{Mab}$. This data further corroborates the functionality of Arr$_{Mab}$ in innate resistance of *M. abscessus* to a wide range of rifamycins, and clarifies KglA is a substrate of Arr$_{Mab}$ (Figure 7.3, E).

In addition, our collaborators in the lab of Prof. Peter Sander at the University of Zurich have gathered MIC data for Rif and KglA with both WT *M. abscessus* and *ΔArr$_{Mab}$ M. abscessus* (unpublished). Their results show *ΔArr$_{Mab}$ M. abscessus* is considerably more sensitive to Rif than the WT strain. Furthermore, there was a marked increase in sensitivity of the *ΔArr$_{Mab}$ M. abscessus* strain to KglA, when compared to WT, further verifying KglA is a substrate for Arr$_{Ms}$.
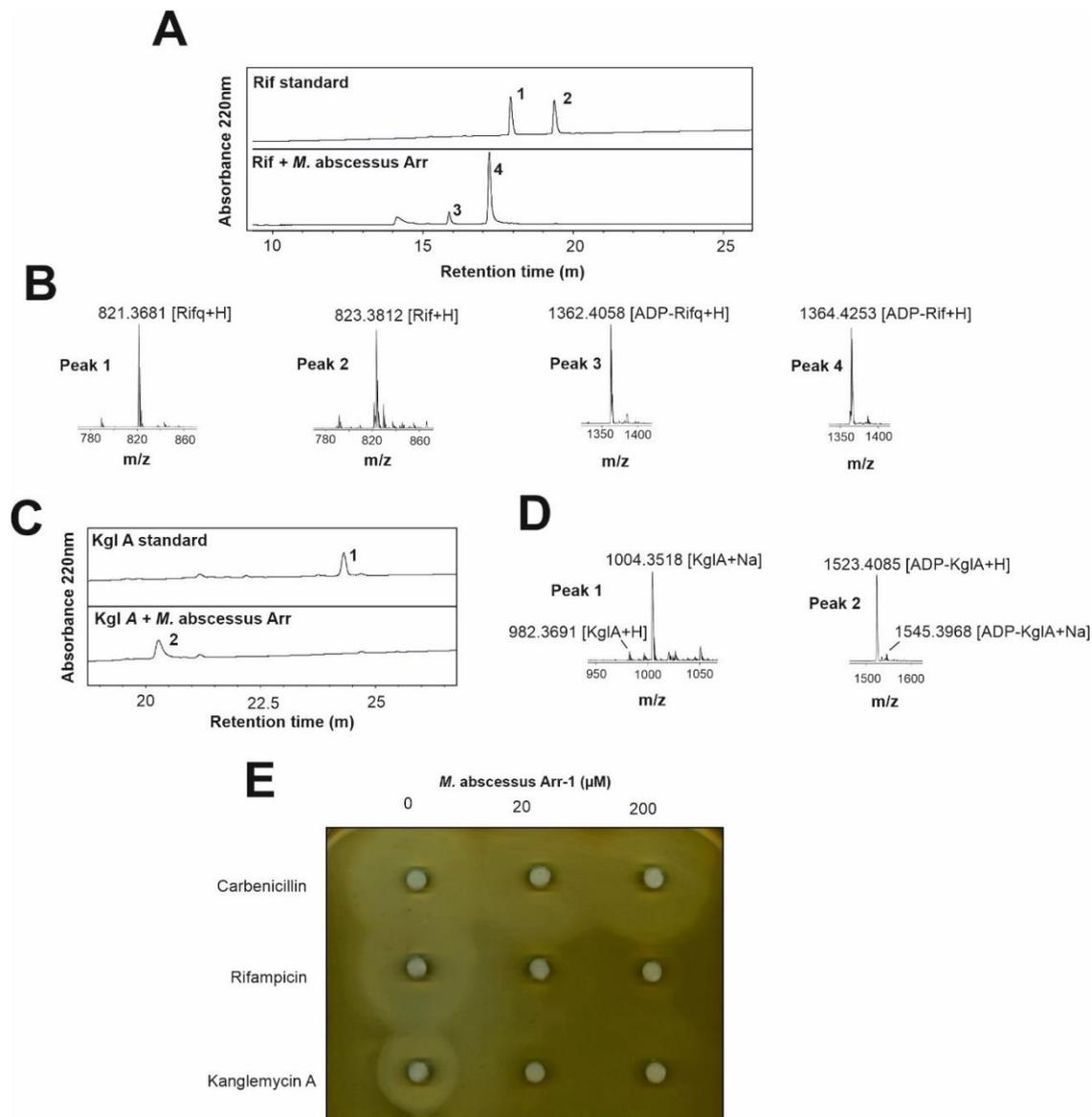
Figure 7.3 *M. abscessus* Arr (Arr$_{Mab}$) inactivates kanglemycin A (KglA) by ADP-ribosylation. Representative reverse phase HPLC traces of *in vitro* inactivation of rifampicin (Rif) by Arr$_{Mab}$ (Upper panel) Trace represents the Rif standard in which Rif is incubated with NAD+ in the absence of Arr$_{Mab}$. (Lower panel) Trace represents Rif incubated with NAD+ for 1h in the presence of Arr$_{Mab}$. Absorbance is at 220nm. (B) Positive ion mass spectrum corresponding to peaks indicated '1', '2', '3' and '4' identified in (A). Ion adducts are as indicated. (C) Representative reverse phase HPLC traces of *in vitro* inactivation of KglA (Rif) with Arr$_{Mab}$. (Upper panel) Trace represents the KglA standard in which Rif is incubated with NAD+ in the absence of Arr$_{Mab}$. (Lower panel) Trace represents KglA incubated with NAD+ for 1h in the presence of Arr$_{Mab}$. Absorbance is at 256nm. (D) Positive ion mass spectrum corresponding to peaks indicated '1' and '2' identified in C. Probable Ion adducts are as indicated. (E) Disk diffusion assay of aliquots of *in vitro* reactions (corresponding to (A) and (C)) with carbenicillin, rifampicin, or kanglemycin A incubated in the presence of increasing concentrations of Arr$_{Mab}$. Paper disks were soaked with aliquots from *in vitro* reactions and were placed on LB agar plates infused with lawn of *S. aureus* RM4220

108

7.2.3 Rifampicin and kanglemycin A binding affinities at Arr$_{Mab}$ and Arr$_{Ms}$

The crystal structure of Arr$_{Ms}$ complexed with Rif provides an explanation for how C25 carbamate derivatives of rifampicin evade inactivation by ADP-ribosylation (Baysarowich et al., 2008). If the binding mode is comparable to Rif, the C25 substituent is expected to sterically clash with the interior of the Arr Rif binding pocket, preventing high affinity association to the protein. Thus far, our data show KglA is not a substrate for Arr$_{Ms}$. We had previously hypothesised the large *ansa* bridge substituents; a 2,2-dimethylsuccinic acid chain at C20 and a β-O-3,4-O,O' methylene digitoxose moiety at C27, may produce a similar steric clash within the antibiotic binding pocket. To ascertain if the additional substituents of KglA preclude binding to Arr$_{Ms}$, we assessed binding affinities of both Rif and KglA at Arr$_{Ms}$ and Arr$_{Mab}$ by microscale thermophoresis (MST). Apparent dissociation constants (K$_d$) were determined by titrating serial dilutions of antibiotic against Arr$_{Ms}$ and Arr$_{Mab}$.

As expected, Rif binds to both Arr$_{Ms}$ and Arr$_{Mab}$, although the compound binds with greater affinity to Arr$_{Ms}$, suggesting the apparent diversity in substrate specificity of Arr$_{Mab}$ comes at a cost of reduced binding affinity (Figure 7.4, A). Conversely, as predicted, KglA binds to Arr$_{Ms}$ with a drastically reduced affinity when compared to Rif, suggesting the bulky 2,2-dimethylsuccinic acid chain at C20 and/or β-O-3,4-O,O' methylene digitoxose at C27 of KglA prevents high affinity binding of the compound (Figure 7.4, B). In contrast, KglA bound to *Arr$_{Mab}$* with an affinity comparable to Rif, further corroborating KglA as a substrate for Arr$_{Mab}$. Indeed, from this data, it is apparent KglA is not a substrate for Arr$_{Ms}$ due to a marked reduction in binding affinity at the enzyme. However, Arr$_{Mab}$ accommodates high affinity binding of KglA, consequently facilitating inactivation by ADP-ribosylation.

In an attempt to identify structural elements that may determine the respective substrate specificities, we overlaid the crystal structure of KglA onto the structure of Arr$_{Ms}$ complexed with Rif using pair fit atomic alignment (Figure 7.4, C, D). If KglA adopts a similar conformation to Rif within the Arr$_{Ms}$ binding pocket, our structure suggests a marked steric clash of the C27 β-O-3,4-O,O' methylene digitoxose moiety with the inner sidewall of the binding pocket. Specifically, the

A Rifampicin

Enzyme	K_d (µg/ml)
● M.smegmatis Arr	5.08 ± 0.3
○ M.abscessus Arr	23.01 ± 4.3

ΔF_{norm}

µg/ml

B Kanglemycin

Enzyme	K_d (µg/ml)
● M. smegmatis Arr	>100
○ M. abscessus Arr	22.25 ± 5.58

ΔF_{norm}

µg/ml

C Rif — Rifampicin binding site

D KglA — Rifampicin binding site — steric clash

E

```
M. smegmatis   6  LKPFEVHESGAYLHGTKADLKVSLRIVPGRESNFGASRIMKEVYIIQDIDAGVWGEELSV  65
M. abscessus   4  GNFFEAHESGAYFHGTKADLAIGELETPGHLSNYSESRVMNIIVEKTIVGGGLARMMGK   63

M. smegmatis  66  SEGRGRIYIVEPEGEIEDDPNVTDKKLPGNPTRSYFTREPVRIVGELTDWEGHSPLQIAA 125
M. abscessus  64  SEGRGHVYIVEPEGTLEDDPNVTDKKFPGNPIHSYRSREPVRIVGEVLDWEGPPPEFMET 123

M. smegmatis 126  MFEGLEDIRRKGLAVIYL 143
M. abscessus 124  FRAGIADIKRKGNAVIYL 141
```

*

F KglA — Rifampicin binding site — NAD+ binding site — steric clash — α1 helix

Figure 7.4 Substrate specificity of *M. smegmatis* and *M. abscessus* Arr. (A) Binding affinities of rifampicin at labelled *M. abscessus* Arr and labelled *M .smegmatis* Arr identified by microscale thermophoresis. (B) Binding affinities of kanglemycin A at labelled *M. abscessus* Arr and labelled *M .smegmatis* Arr identified by microscale thermophoresis. For both (A) and (B), Fnorm (normalized fluorescence = fluorescence after thermophoresis/initial fluorescence) is plotted against antibiotic concentration. Error bars are ± SD. (C) Crystal structure of Rif (green stick model) complexed with *M. smegmatis Arr* (Baysarowich et al., 2008). Arr is shown as semi-transparent surface model. (D) Modelling of KglA (blue stick model, side chains are depicted in yellow) in the Rif binding pocket of *M. smegmatis* Arr (2HW2). Arr is shown as semi-transparent surface model. Projected steric clash with interior of Arr rifampicin binding pocket is indicated. (E) Sequence alignment of Arr *M. smegmatis* and *M. abscessus* with regions of homology indicated. The alpha helix structure (Arr_{Mab} α1) highlighted in (F) is indicated with a red asterisk. (F) Modelling of KglA (blue stick model, side chains are depicted in yellow) in the Rif binding pocket of *M. smegmatis* Arr (2HW2). Arr is shown as gray ribbon model. *M. abscessus* low homology alpha helix (Arr_{Mab} α1) is identified within the structure of *M. smegmatis* Arr, indicated in red ribbon/stick model, and projected steric clash with interior of M. smegmatis Arr rifampicin binding pocket is indicated.

bulky substituent would clash with an alpha helix structure (residues 54-65) that constitutes one half of the Rif binding cleft (Figure 7.4, E, F). This alpha helix, termed α1, is implicated in Rif binding interactions; residue D55 makes polar interactions with O11 of Rif, whilst A56, W59, G60 and L63 form non-polar interactions with the carbon backbone of the *ansa-* bridge (Baysarowich et al., 2008).

To ascertain how this steric clash may be overcome in *Mycobacterium abscessus*, we analysed sequence homology of Arr_{Ms} and Arr_{Mab}. Overall, the amino acid sequences are strikingly similar, exhibiting a 67.4% sequence homology (Figure 7.4, E). However, Arr_{Mab} α1 bears almost no sequence homology with Arr_{Ms}. This lack of homology may underlie structural differences that determine the broader substrate specificity exhibited by Arr_{Mab}. Furthermore, the variation of α1 residues involved in Rif binding may underlie the reduced affinity with which Arr_{Mab} binds Rif, perhaps as a consequence of reduced or altered binding contacts with the respective substrate. Yet, structural analysis of the respective Arr enzymes complexed with KglA and Rif is required in order to confirm the basis of such variations.

7.2.4 ADP-ribosylation renders Rifampicin inactive at RNA polymerases *in vitro*

ADP-ribosylation of Rif by *Arr* enzymes occurs at the hydroxyl group at the C23 position; an essential oxygen functionality in binding of Rif to RNAPs. Indeed, early SAR experimentation demonstrated any alteration of this functionality abolished activity of the compound (Aristoff et al., 2010, Bacchi et al., 1998).
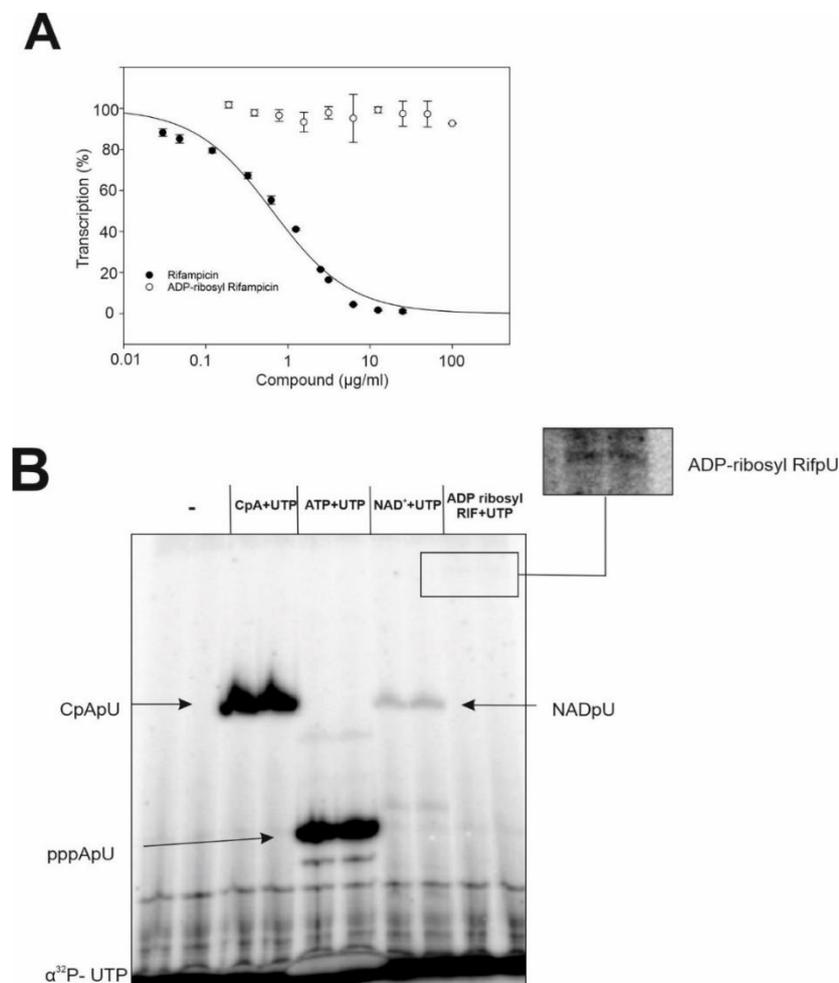
Figure 7.5 ADP-ribosyl rifamycins fail to inhibit RNAP and do not function as an initiating substrate. (A) Quantification of inhibition of *in vitro* transcription by rifampicin and ADP-ribosyl rifampicin. Transcription was performed by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. Reactions were initiated with 5'- non-phosphorylated dinucleotide primer CpA. Quantification is derived from average of [^{32}P]-labelled run-off and terminated transcription products. Values were normalised to quantity of [^{32}P] RNA synthesised in the absence of inhibitor. Error bars are ± SD. (B) A representative gel of *in vitro* abortive transcription performed by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. Reactions were initiated with substrates indicated. Note the negligible transcription initiated by ADP-ribosyl rifampicin.

Despite the assertion that ADP-ribosylation of Rif prevents binding of the inhibitor to RNAP, ADP-ribosyl Rif has never been analysed against RNAP *in vitro*. Consequently, we produced ADP-ribosyl Rif by incubating Rif with purified Arr_{Ms} and isolating the compound with a C8 solid phase extraction cartridge. We then analysed the effects of purified ADP-ribosyl Rif on transcription by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter. Expectedly, ADP-ribosyl Rif was completely unable to inhibit transcription even at high concentrations, showing ADP-ribosylation renders rifampicin ineffectual by inactivating inhibitory activity (Figure 7.5, A).

Interestingly, the structure of ADP ribosyl-Rif possesses an accessible 3' hydroxyl group on the ribosyl moiety that may subsequently act as an initiating substrate for RNAP. Despite, poor binding to RNAP, we assessed if high concentrations of ADP-ribosylated Rif can act as an initiator of abortive transcription by WT *E. coli* RNAP on a linear DNA template containing the T7A1 promoter, and compared abortive synthesis to several other initiating substrates; CpA, ATP and NAD+. Cognate ATP and dinucleotide CpA are efficient initiating substrates of abortive synthesis, whereas transcription initiation by NAD+ is comparatively less efficient (Figure 7.5, B). However, abortive transcription initiating from ADP-ribosyl Rif was virtually undetectable, when compared to CpA, ATP and NAD+. The great steric bulk of ADP ribosyl Rif and its required orientation within the i site of RNAP would likely produce a series of large steric clashes with the interior of RNAP, producing an extremely high energy initiation complex. Indeed, when considering the high concentration of substrate (1mM) used in this instance, under physiological conditions the amount of transcription initiation from ADP- ribosyl Rif would be negligible. Unfortunately, lack of material prevented further reciprocal analysis of ADP-ribosyl KglA *in vitro*. However, ADP- ribosylation would likely prevent the compound binding to RNAP in a manner identical to Rif. Furthermore, the increased bulk of KglA, when compared to Rif, would further reduce the suitability of ADP-ribosyl KglA as an initiating substrate. Indeed, our results indicates ADP-ribosylation of rifamycins cannot subsequently be used as initiating substrates by RNAP.

7.3 Discussion

The communal existence of bacteria generates pressure to gain a competitive edge over neighbouring microorganisms. Consequently, bacteria have evolved a number of mechanisms to counter antibiotics secreted by microbes in their surroundings. Indeed, resistance determinants are found almost ubiquitously amongst bacteria (Crofts et al., 2017). The secondary metabolite precursors of clinical rifamycins are produced by several actinomycetes from highly competitive marine and terrestrial environments. Furthermore, synthetic Rif derivatives have been used to treat a number of clinical infections for over 60 years (Wang et al., 2012, Kim et al., 2006). Consequently, many pathogenic and non-pathogenic bacteria have developed a diverse array of strategies to survive in the presence of rifamycins (see section 1.3.1 for further details).

Rifamycins inhibit transcription by targeting the β-subunit on bacterial RNAP, consequently preventing translocation of the nascent transcript and inhibiting formation of the second or third phosphodiester bond (McClure and Cech, 1978). Resistance to rifamycins is most commonly conferred through point mutations within the Rif binding pocket on the β-subunit (Campbell et al., 2001). However, our work here has shown the growing relevance of rifamycin antibiotic inactivation by ADP-ribosylation in pathogenic Mycobacteria.

ADP-ribosylation of Rif by Arr was first discovered as an intrinsic mechanism of resistance in *Mycobacterium smegmatis*. Arr$_{Ms}$ utilises an NAD+ cosubstrate to catalyse the ADP-ribosylation of Rif at the C23 hydroxyl group, with concurrent expulsion of the nicotinamide moiety from NAD+ (Morisaki et al., 2000) (Figure 7.1). It had been proposed ADP-ribosylation by Arr enzymes inactivates Rif by eliminating the C23 hydroxyl group essential for binding to RNAP. Furthermore, addition of the ADP-ribosyl at C23 orientates the bulky substituent toward the surface of the Rif binding pocket on RNAP, likely precluding binding Rif by severe steric clash (Campbell et al., 2001). Indeed, here we show for the first time ADP-ribosylation of Rif completely abolishes activity of Rif at RNAP *in vitro*, most likely by precluding binding of the compound to RNAP. Furthermore, we show the ADP-ribosyl moiety present on ADP-ribosyl Rif cannot function as an initiating

substrate for RNAP. We believe this mechanism of covalent inactivation is likely to be mutual to all bacterial Arr enzymes.

In addition to Rif, the natural product rifamycin SV, and semi-synthetic Rif derivatives rifaxamin and rifabutin, are also substrates for Arr$_{Ms}$, indicating a broad substrate specificity (Baysarowich et al., 2008). However, the discovery that bulky C25 carbamate Rif derivatives are not subject to ADP- ribosylation by Arr$_{Ms}$ indicated ansamycins with large steric bulk may be able to evade the enzyme, seemingly by no longer associating with the Arr Rif binding pocket (Combrink et al., 2007, Baysarowich et al., 2008). The RNAP inhibitor KglA is a novel ansamycin antibiotic with unique bulky substituents present on the *ansa* bridge; a 2,2-dimethylsuccinic acid chain at C20 and a β-O-3,4-O,O' methylene digitoxose moiety at C27 (Peek et al., 2018, Mosaei et al., 2018) (Figure 7.1). We hypothesised these unique sidechains may prohibit the antibiotic binding to Arr$_{Ms}$, and consequently evade ADP- ribosylation. Indeed, our data show KglA is not a substrate for Arr$_{Ms}$ *in vitro,* as a consequence of reduced binding affinity at the enzyme. Molecular modelling of KglA bound to Arr$_{Ms}$ indicated a large steric clash with the interior of the binding pocket, offering a plausible explanation for the inability of the compound to associate with Arr$_{Ms}$. Considering KglA also retains activity against Rif^R RNAPs and Rif^R bacteria (Peek et al., 2018, Mosaei et al., 2018), the antibiotic may be a promising lead compound with which to target some pathogens possessing Arr enzymes.

However, recently, a putative Arr encoding gene was identified in the highly pathogenic *Mycobacterium abscessus*, one of the most extensively drug-resistant strains of rapidly growing non-tuberculous mycobacteria (NTM) (Rominski et al., 2016). *M. abscessus* is responsible for a wide range of infectious manifestations, including debilitating infections of the skin and respiratory system. A principal challenge in treating infections caused by *M. abscessus* is the bacteria's extraordinary level of innate resistance to a plethora of clinical antibiotics (Luthra et al., 2018). Indeed, deletion of putative Arr$_{Mab}$ increases *M. abscessus* sensitivity to Rif almost 500 fold. Yet, C25 modified rifamycins, which evaded activity at Arr$_{Ms}$, showed equal activity against both WT *M. abscessus,* and an *M. abscessus* mutant lacking *Arr$_{Mab}$,* suggesting the putative Arr$_{Mab}$ may have a differing substrate specificity to Arr$_{Ms}$ (Rominski et al., 2016). Consequently, we

wished to ascertain if, i) The *M. abscessus* genome encodes a viable Arr enzyme, and ii) if so, is the unique ansamycin KglA a substrate for purified Arr$_{Mab}$ *in vitro*. Indeed, our data show the *M. abscessus* genome encodes an active Arr responsible for innate resistance to Rif. Like Arr$_{Ms}$, we show Arr$_{Mab}$ utilises an NAD+ cosubstrate to catalyse the ADP- ribosylation of Rif, presumably at the C23 hydroxyl group, as is the case for Rif (Baysarowich et al., 2008). However, unlike Arr$_{Ms}$, the bulky ansamycin KglA remains a substrate of Arr$_{Mab}$, and is consequently subject to covalent inactivation by ADP-ribosylation. This indicates Arr$_{Mab}$ possesses a divergent substrate specificity to Arr$_{Ms}$, and can seemingly accommodate a broader range of ansamycins with bulky *ansa*-bridge substituents. To examine exactly how Arr$_{Mab}$ can accommodate the binding of larger rifamycin substrates, we performed sequence alignment analysis to identify divergence in the amino acid sequence of the Rif binding pocket of Arr$_{Mab}$ and Arr$_{Ms}$. Indeed, our analysis shows a large portion of the interior of the Rif binding pocket of Arr$_{Mab}$ and Arr$_{Ms}$ lack sequence homology. In particular, an alpha helix, termed α1, implicated in several binding interactions with the Rif *ansa*-bridge, possesses almost no sequence homology between Arr$_{Ms}$ and Arr$_{Mab}$. We propose the differing amino acid constitutions of the interior binding pocket likely underlie structural differences facilitating different substrate specificities. However, direct structural analysis of Arr$_{Mab}$ complexed with KglA is required to ascertain the exact structural basis of ansamycin substrate binding. Furthermore, structural analysis of how different Arr enzymes bind varying substrates is essential for the rational design of novel antibiotics which can evade Arr. Indeed, understanding of how ADP- ribosylation mechanisms vary amongst different bacteria is essential if this abundant resistance mechanism is to be nullified.

Exactly how *M. abscessus* Arr and *M. smegmatis* Arr have diverged to accommodate a different range of ansamycin substrates has not yet been investigated. Nevertheless, competition amongst environmental bacteria places strong evolutionary pressure on production of variant antibiotics capable of avoiding common resistance mechanisms, while synchronously pressurising the development of resistance mechanisms with which to evade these novel variants. Indeed, bulky ansamycin compounds, such as KglA, are produced by actinomycetes that occupy the same natural niches as NTMs, such as *M.*

abscessus (Peek et al., 2018, Honda et al., 2018). Consequently, it is tempting to speculate KglA represents an evolved solution to both spontaneous and innate resistance mechanisms of bacterial competitors (Peek et al., 2018, Mosaei et al., 2018). Subsequently, producers of KglA placed a selective advantage upon competitors capable of surviving in their presence, leading to the development of more diverse Arr enzymes capable of inactivating novel, more complex rifamycins. Indeed, there is now strong genomic evidence that most innate resistance mechanisms seen in clinical pathogens are of environmental origin (Peterson and Kaur, 2018). Consequently, many pathogenic bacteria may already possess a reservoir of resistance determinants to nullify as of yet undiscovered antibiotic variants. Therefore, a comprehensive understanding of resistance mechanisms is essential if we are to capitalise upon future antibiotic discoveries and facilitate the rational development of efficacious antibiotics with which to treat drug resistant infections.

8. Concluding Remarks

Transcription, the first stage of gene expression, is performed by the multi-subunit RNA polymerase (RNAP). The indispensable nature of transcription and sequence divergence from eukaryotic counterparts make bacterial RNAP an excellent target for antibiotics. However, very few clinical antibiotics target RNAP. The growing prevalence of antibiotic resistance amongst pathogenic bacteria demands the identification of novel antibacterial compounds acting through novel molecular mechanisms. Here, we conducted four distinct projects in which we investigated the molecular mechanisms of several previously uncharacterised transcription inhibitors.

Most clinical antibiotics are derived from the natural products of actinomycete bacteria. Our industrial collaborators DemurisTM previously compiled a library of actinomycete bacteria that activate an RNAP reporter strain and therefore may produce novel inhibitors of bacterial transcription. Consequently, we aimed to identify and characterise novel inhibitors of bacterial transcription produced by strains from this particular strain library.

Firstly - our industrial collaborators Demuris^{TM} previously compiled a library of actinomycete bacteria that activate the yvgS RNAP reporter strain and therefore may produce novel inhibitors of bacterial transcription. Subsequently, we aimed to identify and characterise novel inhibitors of bacterial transcription produced by strains from this particular strain library. We utilised *In vitro* transcription assays combined with mass spectroscopic analysis to identify DEM40380 as a producer of Antibiotic A39079S-1, a broad spectrum ansamycin antibiotic with a previously undefined mechanism of action (Boeck, 1985). Here, we have presented data suggesting the compound inhibits transcription through a steric occlusion mechanism typical of rifamycins, likely by targeting the Rif binding pocket on RNAP. The compound possesses the distinctive naphthelenic, 17-mer *ansa* chain structure characteristic of rifamycin antibiotics. However, the activity of antibiotic A39079S-1 sheds new light on the structure-activity relationship of rifamycins. The compound structure indicates a methoxy group at C27, common to most rifamycins, can seemingly be cleaved to a hydroxyl without abolishing activity. Furthermore, the absence of the common C16 methyl group in Antibiotic

A39079S-1 suggests this substituent is not required for the correct conformation of essential oxygen functionalities at C1, C8, C21 and C23. This observation may point to a site on the *ansa-* bridge where small substituents can be introduced successfully to rifamycins.

Secondly - the recently discovered antibiotic ureidothiophene (Urd) was identified within a commercial screen of synthetic compounds in which inhibition of *S. aureus* RNAP was analysed. Here, we have shown the inhibitor targets regulatory sub-region 1.2 of the sigma subunit to prevent melting of the -10 promoter element. The compound consequently prevents formation of the transcription competent open promoter complex. Urd inhibition is dependent upon the integrity of σR1.2, suggesting the compound may directly, or perhaps indirectly (allosterically), interact with this particular σ factor sub-region. σR1.2 has previously been shown to play a vital role in formation of stable open complexes (Wilson and Dombroski, 1997, Baldwin and Dombroski, 2001, Zenkin et al., 2007). Specifically, it is proposed to allosterically regulate recognition of the -10 promoter element by σR2.3. By interacting with σR1.2, we believe Urd interferes with this regulatory signal between σR1.2 and σR2.3, preventing the formation of the open promoter complex.

Thirdly - a prior screening program conducted by our industrial collaboraters Demuris^{TM} had identified the rifamycin type natural product kanglemycin A (KglA) as an inhibitor of rifampicin resistant RNAPs. Here, we show the unique ansa-bridge substituents of the compound act to form new binding contacts with RNAP. We also present data showing KglA inhibits transcription through a unique steric occlusion mechanism, distinct from that of Rif. During *de novo* transcription, KglA inhibits the formation of the first phosphodiester bond, whereas Rif inhibits the formation of the second phosphodiester bond. With our collaborators, the combined biochemical and structural approach has allowed us to discern the basis of this mechanistic difference. The C20 2,2-dimethylsuccinic acid moiety appears to afford an additional steric and electrostatic obstacle to both initiating nucleotide binding and to translocation of the nascent transcript. Furthermore, we have shown KglA also inhibits mycobacterial RNAPs, and consequently retains activity against multi-drug resistant *M. tuberculosis.*

Finally - we investigated ADP- ribosylation as a mechanism of KglA inactivation by *Mycobacterium smegmatis* and *Mycobacterium abscessus* rifampicin ADP-ribosyltransferase (Arr) enzymes. We show KglA is not a substrate of the rifampicin inactivating Arr purified from *Mycobacterium smegmatis*, but remains a substrate of Arr purified from the extensively drug resistant pathogen *Mycobacterium abscessus*. We determined dissociation constants for KglA at both Arr enzymes and discovered KglA is unable to bind to Arr$_{Ms}$, yet is able to bind to Arr$_{Mab}$. Molecular modelling of KglA bound to Arr$_{Ms}$ indicated a large steric clash with the interior of the binding pocket, offering an explanation for the inability of the compound to associate with the enzyme. To identify the basis for the differing substrate specificities between Arr enzymes, we performed sequence alignment analysis to identify divergence in the amino acid sequence of the Rif binding pocket of Arr$_{Mab}$ and Arr$_{Ms}$. Our analysis indicates a large portion of the interior of the Rif binding pocket of Arr$_{Mab}$ and Arr$_{Ms}$ lack sequence homology. Specifically, an alpha helix, termed α1, implicated in several binding interactions with the Rif *ansa*-bridge, possesses almost no sequence homology between Arr$_{Ms}$ and Arr$_{Mab}$. We suggest the differing amino acid constitutions of the interior binding pocket likely underlie structural differences facilitating different substrate specificities. Additionally, we have shown ADP-ribosylation of Rif completely abolishes activity of Rif at RNAP *in vitro*, most likely by precluding binding of the compound to RNAP. Also, we have also shown the ADP-ribosyl moiety present on ADP-ribosyl Rif cannot function as an initiating substrate for RNAP.

By investigating these unique mechanistic processes, we have advanced our understanding of how transcription targeting antibiotics function at RNAP. Furthermore, we have furthered our understanding of the mechanisms utilised by pathogenic bacteria to facilitate resistance to transcription targeting antibiotics.

9. Appendix

T7A1-F 5'-GGTCGACTCTAGAGGATCGCT-3'

T7A1-R 5'-Bio-CGACGTTGTAAAACGACGGCCAGTG-3'

***lac*UV5-F** 5'-CTCACTCATTAGGCACCCCAGGC-3'

***lac*UV5-R** 5'-Bio- CCAGGCGGTGAAGGGCAATCAGC

T7A2-F 5'-CGCTTAAGTCACCTAGAAGGC-3'

T7A2-F 5'-TCGACACCGGGGGAATTCGG-3'

***gal*P1-F** 5'-GGCTAAATTCTTGTGTAAACGATTCCA-3'

***gal*P1-R** 5'- CTCATAATTCGCTCCATTAGGCTTATG-3'

βR143A-F 5'-TGTTATCAACGGTACTGAGGCTGTTATCGTTTCCCAGCTG-3'

βR143A-R 5'-CAGCTGGGAAACGATAACAGCCTCAGTACCGTTGATAACA-3'

βS531L-F 5'-CGCACAAACGTCGTATCTTGGCACTCGGCCC-3'

βS531L-R 5'-GGGCCGAGTGCCAAGATACGACGTTTGTGCG-3'

σ⁷⁰E104Q-F 5'GTACGCATGTACATGCGTCAAATGGGC-3'

σ⁷⁰E104Q-R 5'-CAACGGTGCCCATTTGACGCATGTACATGCGTAC-3'

Arr_Ms-F-NdeI 5'-TAAGCAAAGCTTGTCAGTCATAGATGA-3'

Arr_Ms-R-Hind III 5'-TAAGCAAAGCTTATCCTCACCAACCTC-3'

Arr_Mab-F-NdeI 5'-TAAGCACATATGATGGCGAATCCGCC-3'

Arr_Mab-R-Hind III 5'-TAAGCACATATGATGACGATGCCCAA-3'

Figure S1. List of primers used in this work (5'-3'; Bio, Biotin tag)

Plasmid	Resistance Marker	Characteristic	Reference
pTZ19-T7A1	Kanamycin	P_{T7A1}	(Kashlev et al., 1996)
pTZ19- galP1	Kanamycin	P_{galP1}	(Minakhin and Severinov, 2003)
pT7blue-T7A2	Ampicillin	P_{T7A2}	(Yuzenkova et al., 2011)
pIA581	Ampicillin	P_{T7}–rpoA–rpoB: N_{term}:His$_6$–rpoC	(Svetlov and Artsimovitch, 2015)
pGEMABC	Ampicillin	P_{T7}–rpoA–rpoB– rpoC	(Murakami, 2013)
pACYCDuet-1_Ec_rpoZ	Kanamycin	P_{T7}–rpoZ	(Murakami, 2013)
pET28a-σ^{70}	Kanamycin	P_{T7}-rpoD:N_{term}:His$_6$	(Zenkin et al., 2007)
pET28a-σ$^{70\text{-}E104Q}$	Kanamycin	P_{T7}-rpoD^{E104Q}:N_{term}:His$_6$	This work
pET28a-Arr$_{Ms}$	Kanamycin	P_{T7}-M. smegmatis Arr:N_{term}:His$_6$	This work
pET28a-Arr$_{Mab}$	Kanamycin	P_{T7}-M. abscessus Arr:N_{term}:His$_6$	This work

Figure S2. List of plasmids used in this work

Species/Strain	Genotype	Reference
Bacillus subtilus (yvgS)	*B. subtilis* 168 *amyE::P$_{helD}$-lacZcat* reporter strain	(Hutter et al., 2004b)
Staphylococcus aureus (RN4220)	WT	(Nair et al., 2011)
Staphylococcus epidermidis (ATCC12228)	WT	(MacLea and Trachtenberg, 2017)
Mycobacterium smegmatis (MC^2 155)	WT	(Mohan et al., 2015)
Mycobacterium Abscessus (NCTC 13031)	WT	(Ripoll et al., 2009)
Staphylococcus epidermidis- Mut1 (ATCC12228)	*rpoD^{E105Q}*	This work
Escherichia coli (DH5α)	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs
Escherichia coli (T7 express*)*	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10-- TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrCmrr)114::IS10*	New England Biolabs

Figure S3. List of strains used in this work

10. References

ABDELWAHAB, H., MARTIN DEL CAMPO, J. S., DAI, Y., ADLY, C., EL-SOHAIMY, S. & SOBRADO, P. 2016. Mechanism of Rifampicin Inactivation in Nocardia farcinica. *PLOS ONE,* 11**,** e0162578.

ADELMAN, K., YUZENKOVA, J., LA PORTA, A., ZENKIN, N., LEE, J., LIS, J. T., BORUKHOV, S., WANG, M. D. & SEVERINOV, K. 2004. Molecular mechanism of transcription inhibition by peptide antibiotic Microcin J25. *Mol Cell,* 14**,** 753-62.

ALHADID, Y., CHUNG, S., LERNER, E., TAATJES, D. J., BORUKHOV, S. & WEISS, S. 2017. Studying transcription initiation by RNA polymerase with diffusion-based single-molecule fluorescence. *Protein Science,* 26**,** 1278-1290.

ANANYA, R.-S., MICHAEL, J. B. & ROBERT, L. 2016. Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry,* 85**,** 319-347.

ANDERSEN, S. J., QUAN, S., GOWAN, B. & DABBS, E. R. 1997. Monooxygenase-like sequence of a Rhodococcus equi gene conferring increased resistance to rifampin by inactivating this antibiotic. *Antimicrobial Agents and Chemotherapy,* 41**,** 218-221.

ANDRÉ, E., BASTIDE, L., MICHAUX-CHARACHON, S., GOUBY, A., VILLAIN-GUILLOT, P., LATOUCHE, J., BOUCHET, A., GUALTIÉRI, M. & LEONETTI, J.-P. 2006. Novel synthetic molecules targeting the bacterial RNA polymerase assembly. *Journal of Antimicrobial Chemotherapy,* 57**,** 245-251.

ANDRE, E., BASTIDE, L., VILLAIN-GUILLOT, P., LATOUCHE, J., ROUBY, J. & LEONETTI, J. P. 2004. A multiwell assay to isolate compounds inhibiting the assembly of the prokaryotic RNA polymerase. *Assay Drug Dev Technol,* 2**,** 629-35.

ARHIN, F., BÉLANGER, O., CIBLAT, S., DEHBI, M., DELORME, D., DIETRICH, E., DIXIT, D., LAFONTAINE, Y., LEHOUX, D., LIU, J., MCKAY, G. A., MOECK, G., REDDY, R., ROSE, Y., SRIKUMAR, R., TANAKA, K. S. E., WILLIAMS, D. M., GROS, P., PELLETIER, J., PARR, T. R. & FAR, A. R. 2006. A new class of small molecule RNA polymerase inhibitors with activity against Rifampicin-resistant Staphylococcus aureus1. *Bioorganic & Medicinal Chemistry,* 14**,** 5812-5832.

ARISTOFF, P. A., GARCIA, G. A., KIRCHHOFF, P. D. & HOLLIS SHOWALTER, H. D. 2010. Rifamycins – Obstacles and opportunities. *Tuberculosis,* 90**,** 94-118.

ARLET, G., NADJAR, D., HERRMANN, J. L., DONAY, J. L., LAGRANGE, P. H. & PHILIPPON, A. 2001. Plasmid-mediated rifampin resistance encoded by an arr-2-like gene cassette in Klebsiella pneumoniae producing an ACC-1 class C beta-lactamase. *Antimicrobial agents and chemotherapy,* 45**,** 2971-2972.

ARTSIMOVITCH, I., CHU, C., LYNCH, A. S. & LANDICK, R. 2003. A new class of bacterial RNA polymerase inhibitor affects nucleotide addition. *Science,* 302**,** 650-4.

ARTSIMOVITCH, I. & VASSYLYEV, D. G. 2006. Is It Easy to Stop RNA Polymerase? *Cell Cycle,* 5**,** 399-404.

ARTSIMOVITCH, I., VASSYLYEVA, M. N., SVETLOV, D., SVETLOV, V., PEREDERINA, A., IGARASHI, N., MATSUGAKI, N., WAKATSUKI, S., TAHIROV, T. H. & VASSYLYEV, D. G. 2005. Allosteric Modulation of the RNA Polymerase Catalytic Reaction Is an Essential Component of Transcription Control by Rifamycins. *Cell,* 122**,** 351-363.

BACCHI, A., CARCELLI, M. & PELIZZI, G. 2008. Sampling rifamycin conformational variety by cruising through crystal forms: implications for polymorph screening and for biological models. *New Journal of Chemistry,* 32**,** 1725-1735.

BACCHI, A., PELIZZI, G., NEBULONI, M. & FERRARI, P. 1998. Comprehensive Study on Structure–Activity Relationships of Rifamycins: Discussion of Molecular and Crystal Structure and Spectroscopic and Thermochemical Properties of Rifamycin O. *Journal of Medicinal Chemistry,* 41**,** 2319-2332.

BAE, B., FEKLISTOV, A., LASS-NAPIORKOWSKA, A., LANDICK, R. & DARST, S. A. 2015a. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife,* 4**,** e08504.

BAE, B., NAYAK, D., RAY, A., MUSTAEV, A., LANDICK, R. & DARST, S. A. 2015b. CBR antimicrobials inhibit RNA polymerase via at least two bridge-helix cap-mediated effects on nucleotide addition. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** E4178-E4187.

BALDWIN, N. E. & DOMBROSKI, A. J. 2001. Isolation and characterization of mutations in region 1.2 of Escherichia coliσ70. *Molecular Microbiology,* 42**,** 427-437.

BAR-NAHUM, G., EPSHTEIN, V., RUCKENSTEIN, A. E., RAFIKOV, R., MUSTAEV, A. & NUDLER, E. 2005b. A Ratchet Mechanism of Transcription Elongation and Its Control. *Cell,* 120**,** 183-193.

BARNE, K. A., BOWN, J. A., BUSBY, S. J. & MINCHIN, S. D. 1997. Region 2.5 of the Escherichia coli RNA polymerase σ70subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J,* 16.

BASU, R. S., WARNER, B. A., MOLODTSOV, V., PUPOV, D., ESYUNINA, D., FERNÁNDEZ-TORNERO, C., KULBACHINSKIY, A. & MURAKAMI, K. S. 2014. Structural Basis of Transcription Initiation by Bacterial RNA Polymerase Holoenzyme. *Journal of Biological Chemistry,* 289**,** 24549-24559.

BATADA, N. N., WESTOVER, K. D., BUSHNELL, D. A., LEVITT, M. & KORNBERG, R. D. 2004. Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 17361-17364.

BAUER, A. W., KIRBY, W. M., SHERRIS, J. C. & TURCK, M. 1966. Antibiotic susceptibility testing by a standardized single disk method. *Am J Clin Pathol,* 45**,** 493-6.

BAYRO, M. J., MUKHOPADHYAY, J., SWAPNA, G. V., HUANG, J. Y., MA, L. C., SINEVA, E., DAWSON, P. E., MONTELIONE, G. T. & EBRIGHT, R. H. 2003. Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. *J Am Chem Soc,* 125**,** 12382-3.

BAYSAROWICH, J., KOTEVA, K., HUGHES, D. W., EJIM, L., GRIFFITHS, E., ZHANG, K., JUNOP, M. & WRIGHT, G. D. 2008. Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proceedings of the National Academy of Sciences,* 105**,** 4886-4891.

BELOGUROV, G. A., VASSYLYEVA, M. N., SEVOSTYANOVA, A., APPLEMAN, J. R., XIANG, A. X., LIRA, R., WEBBER, S. E., KLYUYEV, S., NUDLER, E., ARTSIMOVITCH, I. & VASSYLYEV, D. G. 2009. Transcription inactivation through local refolding of the RNA polymerase structure. *Nature,* 457**,** 332-335.

BOCHKAREVA, A. & ZENKIN, N. 2013. The σ 70 region 1.2 regulates promoter escape by unwinding DNA downstream of the transcription start site. *Nucleic Acids Research,* 41**,** 4565-4572.

BOECK, R. D. J. E. K. D. 1985. *Antibiotic A39079 factor S-1 and process for its production*.

BORUKHOV, S. & NUDLER, E. 2003. RNA polymerase holoenzyme: structure, function and biological implications. *Current Opinion in Microbiology,* 6**,** 93-100.

BORUKHOV, S., SAGITOV, V., JOSAITIS, C. A., GOURSE, R. L. & GOLDFARB, A. 1993. Two modes of transcription initiation in vitro at the rrnB P1 promoter of Escherichia coli. *Journal of Biological Chemistry,* 268**,** 23477-82.

BOYACI, H., CHEN, J., JANSEN, R., DARST, S. A. & CAMPBELL, E. A. 2019. Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature,* 565**,** 382-385.

BOYACI, H., CHEN, J., LILIC, M., PALKA, M., MOONEY, R. A., LANDICK, R., DARST, S. A. & CAMPBELL, E. A. 2018. Fidaxomicin jams Mycobacterium tuberculosis RNA polymerase motions needed for initiation via RbpA contacts. *eLife,* 7**,** e34823.

BRAFFMAN, N. R., PISCOTTA, F. J., HAUVER, J., CAMPBELL, E. A., LINK, A. J. & DARST, S. A. 2019. Structural mechanism of transcription inhibition by lasso peptides microcin J25 and capistruin. *Proceedings of the National Academy of Sciences,* 116**,** 1273-1278.

BUC, H. & MCCLURE, W. R. 1985. Kinetics of open complex formation between Escherichia coli RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry,* 24**,** 2712-2723.

BURGESS, R. R. & ANTHONY, L. 2001. How sigma docks to RNA polymerase and what sigma does. *Curr Opin Microbiol,* 4**,** 126-31.

BURGESS, R. R., TRAVERS, A. A., DUNN, J. J. & BAUTZ, E. K. 1969. Factor stimulating transcription by RNA polymerase. *Nature,* 221.

BURTON, Z. F., FEIG, M., GONG, X. Q., ZHANG, C., NEDIALKOV, Y. A. & XIONG, Y. 2005. NTP-driven translocation and regulation of downstream template opening by multi-subunit RNA polymerases. *Biochemistry and Cell Biology,* 83**,** 486-496.

BUURMAN, E. T., FOULK, M. A., GAO, N., LAGANAS, V. A., MCKINNEY, D. C., MOUSTAKAS, D. T., ROSE, J. A., SHAPIRO, A. B. & FLEMING, P. R. 2012. Novel Rapidly Diversifiable Antimicrobial RNA Polymerase Switch Region Inhibitors with Confirmed Mode of Action in Haemophilus influenzae. *Journal of Bacteriology,* 194**,** 5504-5512.

CAMPBELL, E. A., KORZHEVA, N., MUSTAEV, A., MURAKAMI, K., NAIR, S., GOLDFARB, A. & DARST, S. A. 2001. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell,* 104**,** 901-912.

CAMPBELL, E. A., MUZZIN, O., CHLENOV, M., SUN, J. L., OLSON, C. A., WEINMAN, O., TRESTER-ZEDLITZ, M. L. & DARST, S. A. 2002a. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell,* 9.

CAMPBELL, E. A., MUZZIN, O., CHLENOV, M., SUN, J. L., OLSON, C. A., WEINMAN, O., TRESTER-ZEDLITZ, M. L. & DARST, S. A. 2002b. Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. *Molecular Cell,* 9**,** 527-539.

CAMPBELL, E. A., PAVLOVA, O., ZENKIN, N., LEON, F., IRSCHIK, H., JANSEN, R., SEVERINOV, K. & DARST, S. A. 2005. Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *The EMBO journal,* 24**,** 674-682.

CARPOUSIS, A. J. & GRALLA, J. D. 1980. Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry,* 19**,** 3245-3253.

CARPOUSIS, A. J. & GRALLA, J. D. 1985. Interaction of RNA polymerase with lacUV5 promoter DNA during mRNA initiation and elongation: Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation. *Journal of Molecular Biology,* 183**,** 165-177.

CHAKRABORTY, A., WANG, D., EBRIGHT, Y. W., KORLANN, Y., KORTKHONJIA, E., KIM, T., CHOWDHURY, S., WIGNESHWERARAJ, S., IRSCHIK, H., JANSEN, R., NIXON, B. T., KNIGHT, J., WEISS, S. & EBRIGHT, R. H. 2012. Opening and Closing of the Bacterial RNA Polymerase Clamp. *Science,* 337**,** 591-595.

CHEETHAM, G. M. T., JERUZALMI, D. & STEITZ, T. A. 1999. Structural basis for initiation of transcription from an RNA polymerase–promoter complex. *Nature,* 399**,** 80-83.

CHOPRA, I. 2007. Bacterial RNA polymerase: a promising target for the discovery of new antimicrobial agents. *Curr Opin Investig Drugs,* 8**,** 600-7.

CICILIATO, I., CORTI, E., SARUBBI, E., STEFANELLI, S., GASTALDO, L., MONTANINI, N., KURZ, M., LOSI, D., MARINELLI, F. & SELVA, E. 2004. Antibiotics GE23077, novel inhibitors of bacterial RNA polymerase. I. Taxonomy, isolation and characterization. *J Antibiot (Tokyo),* 57**,** 210-7.

COMBRINK, K. D., DENTON, D. A., HARRAN, S., MA, Z., CHAPO, K., YAN, D., BONVENTRE, E., ROCHE, E. D., DOYLE, T. B., ROBERTSON, G. T. & LYNCH, A. S. 2007. New C25 carbamate

rifamycin derivatives are resistant to inactivation by ADP-ribosyl transferases. *Bioorganic & Medicinal Chemistry Letters,* 17**,** 522-526.

CROFTS, T. S., GASPARRINI, A. J. & DANTAS, G. 2017. Next-generation approaches to understand and combat the antibiotic resistome. *Nature Reviews Microbiology,* 15**,** 422-434.

CRUM, G. F., DEVRIES, W. H., EBLE, T. E., LARGE, C. M. & SHELL, J. W. 1955. Streptolydigin, a new antimicrobial antibiotic. II. Isolation and characterization. *Antibiot Annu,* 3**,** 893-6.

DABBS, E. R., YAZAWA, K., TANAKA, Y., MIKAMI, Y., MIYAJI, M., ANDERSEN, S. J., MORISAKI, N., IWASAKI, S., SHIDA, O., TAKAGI, H. & ET AL. 1995. Rifampicin inactivation by Bacillus species. *J Antibiot (Tokyo),* 48**,** 815-9.

DARST, S. A. 2004. New inhibitors targeting bacterial RNA polymerase. *Trends Biochem Sci,* 29**,** 159-60.

DAVIES, J. & DAVIES, D. 2010. Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews,* 74**,** 417-433.

DEBOER, C., DIETZ, A., SAVAGE, G. M. & SILVER, W. S. 1955. Streptolydigin, a new antimicrobial antibiotic. I. Biologic studies of streptolydigin. *Antibiot Annu,* 3**,** 886-92.

DEGEN, D., FENG, Y., ZHANG, Y., EBRIGHT, K. Y., EBRIGHT, Y. W., GIGLIOTTI, M., VAHEDIAN-MOVAHED, H., MANDAL, S., TALAUE, M., CONNELL, N., ARNOLD, E., FENICAL, W. & EBRIGHT, R. H. 2014. Transcription inhibition by the depsipeptide antibiotic salinamide A. *eLife,* 3**,** e02451.

DICKSON, R. C., ABELSON, J., BARNES, W. M. & REZNIKOFF, W. S. 1975. Genetic regulation: the Lac control region. *Science,* 187**,** 27-35.

DRENNAN, A., KRAEMER, M., CAPP, M., GRIES, T., RUFF, E., SHEPPARD, C., WIGNESHWERARAJ, S., ARTSIMOVITCH, I. & RECORD, M. T. 2012. Key Roles of the Downstream Mobile Jaw of Escherichia coli RNA Polymerase in Transcription Initiation. *Biochemistry,* 51**,** 9447-9459.

DUCHI, D., MAZUMDER, A., MALINEN, A. M., EBRIGHT, R. H. & KAPANIDIS, A. N. 2018. The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Res,* 46**,** 7284-7295.

FEKLISTOV, A., BAE, B., HAUVER, J., LASS-NAPIORKOWSKA, A., KALESSE, M., GLAUS, F., ALTMANN, K.-H., HEYDUK, T., LANDICK, R. & DARST, S. A. 2017. RNA polymerase motions during promoter melting. *Science,* 356**,** 863-866.

FEKLISTOV, A. & DARST, S. A. 2009. Promoter recognition by bacterial alternative σ factors: the price of high selectivity? *Genes & Development,* 23**,** 2371-2375.

FEKLISTOV, A. & DARST, SETH A. 2011. Structural Basis for Promoter −10 Element Recognition by the Bacterial RNA Polymerase σ Subunit. *Cell,* 147**,** 1257-1269.

FEKLISTOV, A., MEKLER, V., JIANG, Q., WESTBLADE, L. F., IRSCHIK, H., JANSEN, R., MUSTAEV, A., DARST, S. A. & EBRIGHT, R. H. 2008. Rifamycins do not function by allosteric modulation of binding of Mg2+ to the RNA polymerase active center. *Proc Natl Acad Sci U S A,* 105**,** 14820-5.

FENG, Y., DEGEN, D., WANG, X., GIGLIOTTI, M., LIU, S., ZHANG, Y., DAS, D., MICHALCHUK, T., EBRIGHT, Y. W., TALAUE, M., CONNELL, N. & EBRIGHT, R. H. 2015. Structural Basis of Transcription Inhibition by CBR Hydroxamidines and CBR Pyrazoles. *Structure (London, England : 1993),* 23**,** 1470-1481.

FENG, Y., ZHANG, Y. & EBRIGHT, R. H. 2016. Structural basis of transcription activation. *Science,* 352**,** 1330.

FLOSS, H. G. & YU, T.-W. 2005. RifamycinMode of Action, Resistance, and Biosynthesis. *Chemical Reviews,* 105**,** 621-632.

FRUTH, M., PLAZA, A., HINSBERGER, S., SAHNER, J. H., HAUPENTHAL, J., BISCHOFF, M., JANSEN, R., MÜLLER, R. & HARTMANN, R. W. 2014. Binding Mode Characterization of Novel RNA Polymerase Inhibitors Using a Combined Biochemical and NMR Approach. *ACS Chemical Biology,* 9**,** 2656-2663.

GILL, S. K. & GARCIA, G. A. 2011. Rifamycin inhibition of WT and Rif-resistant Mycobacterium tuberculosis and Escherichia coli RNA polymerases in vitro. *Tuberculosis,* 91**,** 361-369.

GOLDSTEIN, B. P. 2014. Resistance to rifampicin: a review. *J Antibiot (Tokyo),* 67**,** 625-30.

GOURSE, R. L., ROSS, W. & GAAL, T. 2000. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Molecular Microbiology,* 37**,** 687-695.

GRALLA, J. D., CARPOUSIS, A. J. & STEFANO, J. E. 1980. Productive and abortive initiation of transcription in vitro at the lac UV5 promoter. *Biochemistry,* 19**,** 5864-5869.

GUSAROV, I. & NUDLER, E. 1999. The Mechanism of Intrinsic Transcription Termination. *Molecular Cell,* 3**,** 495-504.

HASSAN, H. M., DEGEN, D., JANG, K. H., EBRIGHT, R. H. & FENICAL, W. 2015. Salinamide F, new depsipeptide antibiotic and inhibitor of bacterial RNA polymerase from a marine-derived Streptomyces sp. *J Antibiot (Tokyo),* 68**,** 206-9.

HAUGEN, S. P., BERKMEN, M. B., ROSS, W., GAAL, T., WARD, C. & GOURSE, R. L. 2006. rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. *Cell,* 125**,** 1069-1082.

HAUGEN, S. P., ROSS, W., MANRIQUE, M. & GOURSE, R. L. 2008a. Fine structure of the promoter-sigma region 1.2 interaction. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 3292-3297.

HAWLEY, D. K. & MCCLURE, W. R. 1980. In vitro comparison of initiation properties of bacteriophage lambda wild-type PR and x3 mutant promoters. *Proceedings of the National Academy of Sciences of the United States of America,* 77**,** 6381-6385.

HAWLEY, D. K. & MCCLURE, W. R. 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. *Nucleic Acids Research,* 11**,** 2237-2255.

HEIN, P. P. & LANDICK, R. 2010. The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biology,* 8**,** 141.

HONDA, J. R., VIRDI, R. & CHAN, E. D. 2018. Global Environmental Nontuberculous Mycobacteria and Their Contemporaneous Man-Made and Natural Niches. *Frontiers in microbiology,* 9**,** 2029-2029.

HOUANG, E. T. S., CHU, Y.-W., LO, W.-S., CHU, K.-Y. & CHENG, A. F. B. 2003. Epidemiology of Rifampin ADP-Ribosyltransferase (arr-2) and Metallo-β-Lactamase (bla_{IMP-4}) Gene Cassettes in Class 1 Integrons in Acinetobacter Strains Isolated from Blood Cultures in 1997 to 2000. *Antimicrobial Agents and Chemotherapy,* 47**,** 1382-1390.

HSU, L. M. 2009. Monitoring abortive initiation. *Methods,* 47**,** 25-36.

HSU, L. M., COBB, I. M., OZMORE, J. R., KHOO, M., NAHM, G., XIA, L., BAO, Y. & AHN, C. 2006. Initial Transcribed Sequence Mutations Specifically Affect Promoter Escape Properties. *Biochemistry,* 45**,** 8841-8854.

HSU, L. M., VO, N. V., KANE, C. M. & CHAMBERLIN, M. J. 2003. In Vitro Studies of Transcript Initiation by Escherichia coli RNA Polymerase. 1. RNA Chain Initiation, Abortive Initiation, and Promoter Escape at Three Bacteriophage Promoters. *Biochemistry,* 42**,** 3777-3786.

HUI, J., GORDON, N. & KAJIOKA, R. 1977. Permeability Barrier to Rifampin in Mycobacteria. *Antimicrobial Agents and Chemotherapy,* 11**,** 773-779.

HUTTER, B., FISCHER, C., JACOBI, A., SCHAAB, C. & LOFERER, H. 2004b. Panel of Bacillus subtilis reporter strains indicative of various modes of action. *Antimicrobial agents and chemotherapy,* 48**,** 2588-2594.

IRSCHIK, H., JANSEN, R., GERTH, K., HOFLE, G. & REICHENBACH, H. 1987. The sorangicins, novel and powerful inhibitors of eubacterial RNA polymerase isolated from myxobacteria. *J Antibiot (Tokyo),* 40**,** 7-13.

ISHIHAMA, A. 1992. Role of the RNA polymerase α subunit in transcription activation. *Molecular Microbiology,* 6**,** 3283-3288.

ISHIKAWA, J., YAMASHITA, A., MIKAMI, Y., HOSHINO, Y., KURITA, H., HOTTA, K., SHIBA, T. & HATTORI, M. 2004. The complete genomic sequence of Nocardia farcinica IFM 10152. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 14925-14930.

JACKSON, N., CZAPLEWSKI, L. & PIDDOCK, L. J. V. 2018. Discovery and development of new antibacterial drugs: learning from experience? *Journal of Antimicrobial Chemotherapy,* 73**,** 1452-1459.

JAMIESON, F. B., GUTHRIE, J. L., NEEMUCHWALA, A., LASTOVETSKA, O., MELANO, R. G. & MEHAFFY, C. 2014. Profiling of rpoB mutations and MICs for rifampin and rifabutin in Mycobacterium tuberculosis. *J Clin Microbiol,* 52**,** 2157-62.

JOHNSON, A. P. 2007. Drug evaluation: OPT-80, a narrow-spectrum macrocyclic antibiotic. *Curr Opin Investig Drugs,* 8**,** 168-73.

KAPANIDIS, A. N., MARGEAT, E., HO, S. O., KORTKHONJIA, E., WEISS, S. & EBRIGHT, R. H. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science,* 314**,** 1144-7.

KASHLEV, M., NUDLER, E., SEVERINOV, K., BORUKHOV, S., KOMISSAROVA, N. & GOLDFARB, A. 1996. Histidine-tagged RNA polymerase of Escherichia coliand transcription in solid phase. *Methods Enzymol,* 274.

KIM, T. K., HEWAVITHARANA, A. K., SHAW, P. N. & FUERST, J. A. 2006. Discovery of a New Source of Rifamycin Antibiotics in Marine Sponge Actinobacteria by Phylogenetic Prediction. *Applied and Environmental Microbiology,* 72**,** 2118-2125.

KIREEVA, M. L., OPRON, K., SEIBOLD, S. A., DOMECQ, C., CUKIER, R. I., COULOMBE, B., KASHLEV, M. & BURTON, Z. F. 2012. Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophysics,* 5**,** 11.

KNAPPE, T. A., LINNE, U., ZIRAH, S., REBUFFAT, S., XIE, X. & MARAHIEL, M. A. 2008. Isolation and Structural Characterization of Capistruin, a Lasso Peptide Predicted from the Genome Sequence of Burkholderia thailandensis E264. *Journal of the American Chemical Society,* 130**,** 11446-11454.

KO, J. & HEYDUK, T. 2014. Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse. *Biochemical Journal,* 463**,** 135-144.

KOO, B.-M., RHODIUS, V. A., NONAKA, G., DEHASETH, P. L. & GROSS, C. A. 2009. Reduced capacity of alternative σs to melt promoters ensures stringent promoter recognition. *Genes & Development,* 23**,** 2426-2436.

KORZHEVA, N. & MUSTAEV, A. 2001. Transcription elongation complex: structure and function. *Current Opinion in Microbiology,* 4**,** 119-125.

KORZHEVA, N., MUSTAEV, A., KOZLOV, M., MALHOTRA, A., NIKIFOROV, V., GOLDFARB, A. & DARST, S. A. 2000. A Structural Model of Transcription Elongation. *Science,* 289**,** 619-625.

KORZHEVA, N., MUSTAEV, A., NUDLER, E., NIKIFOROV, V. & GOLDFARB, A. 1998. Mechanistic model of the elongation complex of Escherichia coli RNA polymerase. *Cold Spring Harb Symp Quant Biol,* 63**,** 337-45.

KUZNEDELOV, K., MINAKHIN, L. & SEVERINOV, K. 2003. Preparation and Characterization of Recombinant Thermus aquaticus RNA Polymerase. *Methods in Enzymology.* Academic Press.

KUZNEDELOV, K., SEMENOVA, E., KNAPPE, T. A., MUKHAMEDYAROV, D., SRIVASTAVA, A., CHATTERJEE, S., EBRIGHT, R. H., MARAHIEL, M. A. & SEVERINOV, K. 2011. The Antibacterial Threaded-lasso Peptide Capistruin Inhibits Bacterial RNA Polymerase. *Journal of Molecular Biology,* 412**,** 842-848.

LANDICK, R. 2005. NTP-entry routes in multi-subunit RNA polymerases. *Trends in Biochemical Sciences,* 30**,** 651-654.

LEWIS, C., NIKITAS, C. T., SCHWARTZ, D. F. & WILKINS, J. R. 1955. Streptolydigin, a new antimicrobial antibiotic. III. In vitro and in vivo laboratory studies. *Antibiot Annu,* 3**,** 897-902.

LI, X.-Y. & MCCLURE, W. R. 1998. Characterization of the Closed Complex Intermediate Formed during Transcription Initiation by Escherichia coli RNA Polymerase. *Journal of Biological Chemistry,* 273**,** 23549-23557.

LIN, W., DAS, K., DEGEN, D., MAZUMDER, A., DUCHI, D., WANG, D., EBRIGHT, Y. W., EBRIGHT, R. Y., SINEVA, E., GIGLIOTTI, M., SRIVASTAVA, A., MANDAL, S., JIANG, Y., LIU, Y., YIN, R., ZHANG, Z., ENG, E., THOMAS, D., DONADIO, S., ZHANG, H., ZHANG, C., KAPANIDIS, A. & EBRIGHT, R. H. 2017a. Structural basis of transcription inhibition by fidaxomicin (lipiarmycin A3). *bioRxiv.*

LIN, W., MANDAL, S., DEGEN, D., LIU, Y., EBRIGHT, Y. W., LI, S., FENG, Y., ZHANG, Y., MANDAL, S., JIANG, Y., LIU, S., GIGLIOTTI, M., TALAUE, M., CONNELL, N., DAS, K., ARNOLD, E. & EBRIGHT, R. H. 2017b. Structural Basis of Mycobacterium tuberculosis Transcription and Transcription Inhibition. *Mol Cell,* 66**,** 169-179.e8.

LIU, L.-K., ABDELWAHAB, H., MARTIN DEL CAMPO, J. S., MEHRA-CHAUDHARY, R., SOBRADO, P. & TANNER, J. J. 2016. The Structure of the Antibiotic Deactivating, N-hydroxylating Rifampicin Monooxygenase. *Journal of Biological Chemistry,* 291**,** 21553-21562.

LUTHRA, S., ROMINSKI, A. & SANDER, P. 2018. The Role of Antibiotic-Target-Modifying and Antibiotic-Modifying Enzymes in Mycobacterium abscessus Drug Resistance. *Frontiers in Microbiology,* 9.

MA, C., YANG, X. & LEWIS, P. J. 2016. Bacterial Transcription as a Target for Antibacterial Drug Development. *Microbiology and Molecular Biology Reviews,* 80**,** 139-160.

MACLEA, K. S. & TRACHTENBERG, A. M. 2017. Complete Genome Sequence of Staphylococcus epidermidis ATCC 12228 Chromosome and Plasmids, Generated by Long-Read Sequencing. *Genome Announcements,* 5**,** e00954-17.

MAFFIOLI, S. I., ZHANG, Y., DEGEN, D., CARZANIGA, T., DEL GATTO, G., SERINA, S., MONCIARDINI, P., MAZZETTI, C., GUGLIERAME, P., CANDIANI, G., CHIRIAC, A. I., FACCHETTI, G., KALTOFEN, P., SAHL, H. G., DEHO, G., DONADIO, S. & EBRIGHT, R. H. 2017. Antibacterial Nucleoside-Analog Inhibitor of Bacterial RNA Polymerase. *Cell,* 169**,** 1240-1248.e23.

MALINEN, A. M., NANDYMAZUMDAR, M., TURTOLA, M., MALMI, H., GROCHOLSKI, T., ARTSIMOVITCH, I. & BELOGUROV, G. A. 2014. CBR antimicrobials alter coupling between the bridge helix and the β subunit in RNA polymerase. *Nature Communications,* 5**,** 3408.

MATHEW, R. & CHATTERJI, D. 2006. The evolving story of the omega subunit of bacterial RNA polymerase. *Trends in Microbiology,* 14**,** 450-455.

MAZUMDER, A. & KAPANIDIS, A. N. 2019b. Recent Advances in Understanding σ70-Dependent Transcription Initiation Mechanisms. *Journal of Molecular Biology,* 431**,** 3947-3959.

MAZUMDER, A., LIN, M., KAPANIDIS, A. N. & EBRIGHT, R. H. 2019. Closing and opening of the RNA polymerase trigger loop. *bioRxiv***,** 817080.

MCCLURE, W. R. 1980. On the mechanism of streptolydigin inhibition of Escherichia coli RNA polymerase. *J Biol Chem,* 255**,** 1610-6.

MCCLURE, W. R. & CECH, C. L. 1978. On the mechanism of rifampicin inhibition of RNA synthesis. *Journal of Biological Chemistry,* 253**,** 8949-8956.

MEJIA, Y. X., NUDLER, E. & BUSTAMANTE, C. 2015. Trigger loop folding determines transcription rate of Escherichia coli's RNA polymerase. *Proceedings of the National Academy of Sciences,* 112**,** 743-748.

MEKLER, V., MINAKHIN, L., BORUKHOV, S., MUSTAEV, A. & SEVERINOV, K. 2014. Coupling of Downstream RNA Polymerase–Promoter Interactions with Formation of Catalytically Competent Transcription Initiation Complex. *Journal of Molecular Biology,* 426**,** 3973-3984.

131

MINAKHIN, L. & SEVERINOV, K. 2003. On the Role of the Escherichia coli RNA Polymerase σ70 Region 4.2 and α-Subunit C-terminal Domains in Promoter Complex Formation on the Extended –10 galP1 Promoter. *Journal of Biological Chemistry,* 278**,** 29710-29718.

MOHAN, A., PADIADPU, J., BALONI, P. & CHANDRA, N. 2015. Complete Genome Sequences of a Mycobacterium smegmatis Laboratory Strain (MC² 155) and Isoniazid-Resistant (4XR1/R2) Mutant Strains. *Genome Announcements,* 3**,** e01520-14.

MOLODTSOV, V., FLEMING, P. R., EYERMANN, C. J., FERGUSON, A. D., FOULK, M. A., MCKINNEY, D. C., MASSE, C. E., BUURMAN, E. T. & MURAKAMI, K. S. 2015. X-ray crystal structures of Escherichia coli RNA polymerase with switch region binding inhibitors enable rational design of squaramides with an improved fraction unbound to human plasma protein. *Journal of medicinal chemistry,* 58**,** 3156-3171.

MOLODTSOV, V., SCHARF, N. T., STEFAN, M. A., GARCIA, G. A. & MURAKAMI, K. S. 2017b. Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in Mycobacterium tuberculosis. *Molecular Microbiology,* 103**,** 1034-1045.

MOORE, B. S., TRISCHMAN, J. A., SENG, D., KHO, D., JENSEN, P. R. & FENICAL, W. 1999. Salinamides, Antiinflammatory Depsipeptides from a Marine Streptomycete. *The Journal of Organic Chemistry,* 64**,** 1145-1150.

MORICHAUD, Z., CHALOIN, L. & BRODOLIN, K. 2016. Regions 1.2 and 3.2 of the RNA Polymerase σ Subunit Promote DNA Melting and Attenuate Action of the Antibiotic Lipiarmycin. *Journal of Molecular Biology,* 428**,** 463-476.

MORISAKI, N., HASHIMOTO, Y., FURIHATA, K., IMAI, T., WATANABE, K., MIKAMI, Y., YAZAWA, K., ANDO, A., NAGATA, Y. & DABBS, E. R. 2000. Structures of ADP-ribosylated rifampicin and its metabolite: intermediates of rifampicin-ribosylation by Mycobacterium smegmatis DSM43756. *J Antibiot (Tokyo),* 53**,** 269-75.

MOSAEI, H. & HARBOTTLE, J. 2019. Mechanisms of antibiotics inhibiting bacterial RNA polymerase. *Biochemical Society Transactions,* 47**,** 339-350.

MOSAEI, H., MOLODTSOV, V., KEPPLINGER, B., HARBOTTLE, J., MOON, C. W., JEEVES, R. E., CECCARONI, L., SHIN, Y., MORTON-LAING, S., MARRS, E. C. L., WILLS, C., CLEGG, W., YUZENKOVA, Y., PERRY, J. D., BACON, J., ERRINGTON, J., ALLENBY, N. E. E., HALL, M. J., MURAKAMI, K. S. & ZENKIN, N. 2018. Mode of Action of Kanglemycin A, an Ansamycin Natural Product that Is Active against Rifampicin-Resistant Mycobacterium tuberculosis. *Mol Cell,* 72**,** 263-274.e5.

MUKHERJEE, R. & CHATTERJI, D. 2008. Stationary phase induced alterations in mycobacterial RNA polymerase assembly: A cue to its phenotypic resistance towards rifampicin. *Biochemical and Biophysical Research Communications,* 369**,** 899-904.

MUKHOPADHYAY, J., DAS, K., ISMAIL, S., KOPPSTEIN, D., JANG, M., HUDSON, B., SARAFIANOS, S., TUSKE, S., PATEL, J., JANSEN, R., IRSCHIK, H., ARNOLD, E. & EBRIGHT, R. H. 2008. The RNA polymerase "switch region" is a target for inhibitors. *Cell,* 135**,** 295-307.

MUKHOPADHYAY, J., SINEVA, E., KNIGHT, J., LEVY, R. M. & EBRIGHT, R. H. 2004. Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Molecular cell,* 14**,** 739-751.

MURAKAMI, K. S. 2013. X-ray Crystal Structure of Escherichia coli RNA Polymerase σ70 Holoenzyme. *Journal of Biological Chemistry,* 288**,** 9126-9134.

MURAKAMI, K. S. 2015. Structural biology of bacterial RNA polymerase. *Biomolecules,* 5**,** 848-864.

MURAKAMI, K. S. & DARST, S. A. 2003. Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol,* 13**,** 31-9.

MURAKAMI, K. S., MASUDA, S., CAMPBELL, E. A., MUZZIN, O. & DARST, S. A. 2002a. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science,* 296.

MURAKAMI, K. S., MASUDA, S. & DARST, S. A. 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science,* 296.

MUSTAEV, A., ROBERTS, J. & GOTTESMAN, M. 2017. Transcription elongation. *Transcription,* 8**,** 150-161.

NAAS, T., MIKAMI, Y., IMAI, T., POIREL, L. & NORDMANN, P. 2001. Characterization of In53, a Class 1 Plasmid- and Composite Transposon-Located Integron of Escherichia coli Which Carries an Unusual Array of Gene Cassettes. *Journal of Bacteriology,* 183**,** 235-249.

NAIR, D., MEMMI, G., HERNANDEZ, D., BARD, J., BEAUME, M., GILL, S., FRANCOIS, P. & CHEUNG, A. L. 2011. Whole-Genome Sequencing of Staphylococcus aureus Strain RN4220, a Key Laboratory Strain Used in Virulence Research, Identifies Mutations That Affect Not Only Virulence Factors but Also the Fitness of the Strain. *Journal of Bacteriology,* 193**,** 2332-2335.

NUDLER, E. 2009. RNA polymerase active center: the molecular engine of transcription. *Annu Rev Biochem,* 78**,** 335-61.

NUDLER, E., AVETISSOVA, E., MARKOVTSOV, V. & GOLDFARB, A. 1996. Transcription Processivity: Protein-DNA Interactions Holding Together the Elongation Complex. *Science,* 273**,** 211-217.

NUDLER, E., MUSTAEV, A., GOLDFARB, A. & LUKHTANOV, E. 1997. The RNA–DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase. *Cell,* 89**,** 33-41.

PAGET, M. S. 2015. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules,* 5**,** 1245-1265.

PAGET, M. S. B. & HELMANN, J. D. 2003. The σ70family of sigma factors. *Genome Biology,* 4**,** 203.

PALLABI, M., GAIRIKA, G., MD, H. & RANJAN, S. 2017. Rho Protein: Roles and Mechanisms. *Annual Review of Microbiology,* 71**,** 687-709.

PEEK, J., LILIC, M., MONTIEL, D., MILSHTEYN, A., WOODWORTH, I., BIGGINS, J. B., TERNEI, M. A., CALLE, P. Y., DANZIGER, M., WARRIER, T., SAITO, K., BRAFFMAN, N., FAY, A., GLICKMAN, M. S., DARST, S. A., CAMPBELL, E. A. & BRADY, S. F. 2018. Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nature Communications,* 9**,** 4147.

PÉRIGAUD, C., GOSSELIN, G. & IMBACH, J. L. 1992. Nucleoside Analogues as Chemotherapeutic Agents: A Review. *Nucleosides and Nucleotides,* 11**,** 903-945.

PETERSON, E. & KAUR, P. 2018. Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Frontiers in Microbiology,* 9.

REVYAKIN, A., LIU, C., EBRIGHT, R. H. & STRICK, T. R. 2006. Abortive Initiation and Productive Initiation by RNA Polymerase Involve DNA Scrunching. *Science,* 314**,** 1139-1143.

RIPOLL, F., PASEK, S., SCHENOWITZ, C., DOSSAT, C., BARBE, V., ROTTMAN, M., MACHERAS, E., HEYM, B., HERRMANN, J.-L., DAFFÉ, M., BROSCH, R., RISLER, J.-L. & GAILLARD, J.-L. 2009. Non Mycobacterial Virulence Genes in the Genome of the Emerging Pathogen Mycobacterium abscessus. *PLOS ONE,* 4**,** e5660.

ROBERTS, J. W. & ROBERTS, C. W. 1996. Base-Specific Recognition of the Nontemplate Strand of Promoter DNA by E. coli RNA Polymerase. *Cell,* 86**,** 495-501.

ROBERTSEN, H. L. & MUSIOL-KROLL, E. M. 2019. Actinomycete-Derived Polyketides as a Source of Antibiotics and Lead Structures for the Development of New Antimicrobial Drugs. *Antibiotics (Basel),* 8.

ROE, J.-H., BURGESS, R. R. & RECORD, M. T. 1984. Kinetics and mechanism of the interaction of Escherichia coli RNA polymerase with the λPR promoter. *Journal of Molecular Biology,* 176**,** 495-522.

ROMINSKI, A., RODITSCHEFF, A., SELCHOW, P., BÖTTGER, E. C. & SANDER, P. 2016. Intrinsic rifamycin resistance of Mycobacterium abscessus is mediated by ADP-ribosyltransferase MAB_0591. *Journal of Antimicrobial Chemotherapy,* 72**,** 376-384.

ROMMELE, G., WIRZ, G., SOLF, R., VOSBECK, K., GRUNER, J. & WEHRLI, W. 1990. Resistance of Escherichia coli to rifampicin and sorangicin A--a comparison. *J Antibiot (Tokyo),* 43**,** 88-91.

ROSENQVIST, P., PALMU, K., PRAJAPATI, R. K., YAMADA, K., NIEMI, J., BELOGUROV, G. A., METSÄ-KETELÄ, M. & VIRTA, P. 2019. Characterization of C-nucleoside Antimicrobials from Streptomyces albus DSM 40763: Strepturidin is Pseudouridimycin. *Scientific Reports,* 9**,** 8935.

ROSS, W., GOSINK, K. K., SALOMON, J., IGARASHI, K., ZOU, C., ISHIHAMA, A., SEVERINOV, K. & GOURSE, R. L. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science,* 262**,** 1407-1413.

RUFF, E. F., RECORD, M. T., JR. & ARTSIMOVITCH, I. 2015. Initial events in bacterial transcription initiation. *Biomolecules,* 5**,** 1035-1062.

SALOMÓN, R. A. & FARÍAS, R. N. 1992. Microcin 25, a novel antimicrobial peptide produced by Escherichia coli. *Journal of Bacteriology,* 174**,** 7428-7435.

SANDGREN, A., STRONG, M., MUTHUKRISHNAN, P., WEINER, B. K., CHURCH, G. M. & MURRAY, M. B. 2009. Tuberculosis Drug Resistance Mutation Database. *PLOS Medicine,* 6**,** e1000002.

SARUBBI, E., MONTI, F., CORTI, E., MIELE, A. & SELVA, E. 2004. Mode of action of the microbial metabolite GE23077, a novel potent and selective inhibitor of bacterial RNA polymerase. *Eur J Biochem,* 271**,** 3146-54.

SCHÄBERLE, T. F., LOHR, F., SCHMITZ, A. & KÖNIG, G. M. 2014. Antibiotics from myxobacteria. *Natural Product Reports,* 31**,** 953-972.

SENSI, P. 1983. History of the development of rifampin. *Rev Infect Dis,* 5 Suppl 3**,** S402-6.

SENSI, P., GRECO, A. M. & BALLOTTA, R. 1959. Rifomycin. I. Isolation and properties of rifomycin B and rifomycin complex. *Antibiot Annu,* 7**,** 262-70.

SENSI, P., TIMBAL, M. T. & MAFFII, G. 1960. Rifomycin IX. Two new antibiotics of rifomycin family: rifomycin S and rifomycin SV. Preliminary report. *Experientia,* 16**,** 412.

SEVERINOV, K., MARKOV, D., SEVERINOVA, E., NIKIFOROV, V., LANDICK, R., DARST, S. A. & GOLDFARB, A. 1995. Streptolydigin-resistant Mutants in an Evolutionarily Conserved Region of the β′ Subunit of Escherichia coli RNA Polymerase. *Journal of Biological Chemistry,* 270**,** 23926-23929.

SHIMADA, T., YAMAZAKI, Y., TANAKA, K. & ISHIHAMA, A. 2014. The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of Escherichia coli. *PLOS ONE,* 9**,** e90447.

SHIN, J. H., EOM, H., SONG, W. J. & RHO, M. 2018. Integrative metagenomic and biochemical studies on rifamycin ADP-ribosyltransferases discovered in the sediment microbiome. *Scientific Reports,* 8**,** 12143.

SOSUNOV, V., SOSUNOVA, E., MUSTAEV, A., BASS, I., NIKIFOROV, V. & GOLDFARB, A. 2003. Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *The EMBO Journal,* 22**,** 2234-2244.

SOSUNOV, V., ZOROV, S., SOSUNOVA, E., NIKOLAEV, A., ZAKEYEVA, I., BASS, I., GOLDFARB, A., NIKIFOROV, V., SEVERINOV, K. & MUSTAEV, A. 2005. The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Res,* 33.

SPANOGIANNOPOULOS, P., THAKER, M., KOTEVA, K., WAGLECHNER, N. & WRIGHT, G. D. 2012. Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrobial agents and chemotherapy,* 56**,** 5061-5069.

SRIVASTAVA, A., TALAUE, M., LIU, S., DEGEN, D., EBRIGHT, R. Y., SINEVA, E., CHAKRABORTY, A., DRUZHININ, S. Y., CHATTERJEE, S., MUKHOPADHYAY, J., EBRIGHT, Y. W., ZOZULA, A., SHEN, J., SENGUPTA, S., NIEDFELDT, R. R., XIN, C., KANEKO, T., IRSCHIK, H., JANSEN, R., DONADIO, S., CONNELL, N. & EBRIGHT, R. H. 2011. NEW TARGET FOR INHIBITION OF BACTERIAL RNA POLYMERASE: "SWITCH REGION". *Current opinion in microbiology,* 14**,** 532-543.

STOGIOS, P. J., COX, G., SPANOGIANNOPOULOS, P., PILLON, M. C., WAGLECHNER, N., SKARINA, T., KOTEVA, K., GUARNÉ, A., SAVCHENKO, A. & WRIGHT, G. D. 2016. Rifampin phosphotransferase is an unusual antibiotic resistance kinase. *Nature Communications,* 7**,** 11343.

STRANEY, D. C. & CROTHERS, D. M. 1985. Intermediates in transcription initiation from the E. coli lac UV5 promoter. *Cell,* 43**,** 449-459.

SVETLOV, V. & ARTSIMOVITCH, I. 2015. Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol Biol,* 1276**,** 13-29.

SZUMOWSKI, J. D., ADAMS, K. N., EDELSTEIN, P. H. & RAMAKRISHNAN, L. 2013. Antimicrobial efflux pumps and Mycobacterium tuberculosis drug tolerance: evolutionary considerations. *Curr Top Microbiol Immunol,* 374**,** 81-108.

TANAKA, Y., YAZAWA, K., DABBS, E. R., NISHIKAWA, K., KOMAKI, H., MIKAMI, Y., MIYAJI, M., MORISAKI, N. & IWASAKI, S. 1996. Different Rifampicin Inactivation Mechanisms in Nocardia and Related Taxa. *Microbiology and Immunology,* 40**,** 1-4.

TEMIAKOV, D., ZENKIN, N., VASSYLYEVA, M. N., PEREDERINA, A., TAHIROV, T. H., KASHKINA, E., SAVKINA, M., ZOROV, S., NIKIFOROV, V., IGARASHI, N., MATSUGAKI, N., WAKATSUKI, S., SEVERINOV, K. & VASSYLYEV, D. G. 2005a. Structural Basis of Transcription Inhibition by Antibiotic Streptolydigin. *Molecular Cell,* 19**,** 655-666.

TRIBUDDHARAT, C. & FENNEWALD, M. 1999. Integron-Mediated Rifampin Resistance inPseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy,* 43**,** 960-962.

TRISCHMAN, J. A., TAPIOLAS, D. M., JENSEN, P. R., DWIGHT, R., FENICAL, W., MCKEE, T. C., IRELAND, C. M., STOUT, T. J. & CLARDY, J. 1994. Salinamides A and B: anti-inflammatory depsipeptides from a marine streptomycete. *Journal of the American Chemical Society,* 116**,** 757-758.

TUPIN, A., GUALTIERI, M., LEONETTI, J.-P. & BRODOLIN, K. 2010a. The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site. *The EMBO Journal,* 29**,** 2527-2537.

TUPIN, A., GUALTIERI, M., ROQUET-BANÈRES, F., MORICHAUD, Z., BRODOLIN, K. & LEONETTI, J.-P. 2010b. Resistance to rifampicin: at the crossroads between ecological, genomic and medical concerns. *International Journal of Antimicrobial Agents,* 35**,** 519-523.

TUSKE, S., SARAFIANOS, S. G., WANG, X., HUDSON, B., SINEVA, E., MUKHOPADHYAY, J., BIRKTOFT, J. J., LEROY, O., ISMAIL, S., CLARK, A. D., JR., DHARIA, C., NAPOLI, A., LAPTENKO, O., LEE, J., BORUKHOV, S., EBRIGHT, R. H. & ARNOLD, E. 2005. Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell,* 122**,** 541-52.

VASSYLYEV, D. G., SEKINE, S., LAPTENKO, O., LEE, J., VASSYLYEVA, M. N., BORUKHOV, S. & YOKOYAMA, S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature,* 417.

VASSYLYEV, D. G., VASSYLYEVA, M. N., PEREDERINA, A., TAHIROV, T. H. & ARTSIMOVITCH, I. 2007a. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature,* 448**,** 157-162.

VASSYLYEV, D. G., VASSYLYEVA, M. N., ZHANG, J., PALANGAT, M., ARTSIMOVITCH, I. & LANDICK, R. 2007b. Structural basis for substrate loading in bacterial RNA polymerase. *Nature,* 448.

VIGLIOTTA, G., TREDICI, S. M., DAMIANO, F., MONTINARO, M. R., PULIMENO, R., DI SUMMA, R., MASSARDO, D. R., GNONI, G. V. & ALIFANO, P. 2005. Natural merodiploidy involving duplicated rpoB alleles affects secondary metabolism in a producer actinomycete. *Molecular Microbiology,* 55**,** 396-412.

VILLAIN-GUILLOT, P., BASTIDE, L., GUALTIERI, M. & LEONETTI, J.-P. 2007a. Progress in targeting bacterial transcription. *Drug Discovery Today,* 12**,** 200-208.

VILLAIN-GUILLOT, P., GUALTIERI, M., BASTIDE, L. & LEONETTI, J.-P. 2007b. In vitro activities of different inhibitors of bacterial transcription against Staphylococcus epidermidis biofilm. *Antimicrobial agents and chemotherapy,* 51**,** 3117-3121.

VO, N. V., HSU, L. M., KANE, C. M. & CHAMBERLIN, M. J. 2003. In Vitro Studies of Transcript Initiation by Escherichia coli RNA Polymerase. 3. Influences of Individual DNA Elements within the Promoter Recognition Region on Abortive Initiation and Promoter Escape. *Biochemistry,* 42**,** 3798-3811.

WALTER, G., ZILLIG, W., PALM, P. & FUCHS, E. 1967. Initiation of DNA-Dependent RNA Synthesis and the Effect of Heparin on RNA Polymerase. *European Journal of Biochemistry,* 3**,** 194-201.

WANG, D., BUSHNELL, D., WESTOVER, K., KAPLAN, C. & KORNBERG, R. 2006. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell,* 127.

WANG, G., ZHANG, H., SUN, G., WU, L., ZHANG, J. & WANG, Y. 2012. A new method for rapid identification of ansamycin compounds by inactivating KLM gene clusters in potential ansamycin-producing actinomyces. *Journal of Applied Microbiology,* 112**,** 353-362.

WANG, N. J., FU, Y., YAN, G. H., BAO, G. H., XU, C. F. & HE, C. H. 1988. Isolation and structure of a new ansamycin antibiotic kanglemycin A from a Nocardia. *J Antibiot (Tokyo),* 41**,** 264-7.

WEINZIERL, R. O. J. 2010. The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. *BMC Biology,* 8**,** 134.

WIEGAND, I., HILPERT, K. & HANCOCK, R. E. 2008. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc,* 3**,** 163-75.

WILSON, C. & DOMBROSKI, A. J. 1997. Region 1 of σ70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase 11Edited by M. Gottesman. *Journal of Molecular Biology,* 267**,** 60-74.

WILSON, K. A., KALKUM, M., OTTESEN, J., YUZENKOVA, J., CHAIT, B. T., LANDICK, R., MUIR, T., SEVERINOV, K. & DARST, S. A. 2003. Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. *J Am Chem Soc,* 125**,** 12475-83.

WOHLLEBEN, W., MAST, Y., STEGMANN, E. & ZIEMERT, N. 2016. Antibiotic drug discovery. *Microbial Biotechnology,* 9**,** 541-548.

XU, M., ZHOU, Y. N., GOLDSTEIN, B. P. & JIN, D. J. 2005. Cross-resistance of Escherichia coli RNA polymerases conferring rifampin resistance to different antibiotics. *Journal of bacteriology,* 187**,** 2783-2792.

YANG, X. & PRICE, C. W. 1995. Streptolydigin Resistance Can Be Conferred by Alterations to Either the β or β' Subunits of Bacillus subtilis RNA Polymerase. *Journal of Biological Chemistry,* 270**,** 23930-23933.

YUZENKOVA, J., DELGADO, M., NECHAEV, S., SAVALIA, D., EPSHTEIN, V., ARTSIMOVITCH, I., MOONEY, R. A., LANDICK, R., FARIAS, R. N., SALOMON, R. & SEVERINOV, K. 2002. Mutations of bacterial RNA polymerase leading to resistance to microcin j25. *J Biol Chem,* 277.

YUZENKOVA, Y., BOCHKAREVA, A., TADIGOTLA, V. R., ROGHANIAN, M., ZOROV, S., SEVERINOV, K. & ZENKIN, N. 2010. Stepwise mechanism for transcription fidelity. *BMC Biology,* 8**,** 54.

YUZENKOVA, Y., ROGHANIAN, M., BOCHKAREVA, A. & ZENKIN, N. 2013. Tagetitoxin inhibits transcription by stabilizing pre-translocated state of the elongation complex. *Nucleic acids research,* 41**,** 9257-9265.

YUZENKOVA, Y., TADIGOTLA, V. R., SEVERINOV, K. & ZENKIN, N. 2011. A new basal promoter element recognized by RNA polymerase core enzyme. *The EMBO journal,* 30**,** 3766-3775.

ZAYCHIKOV, E., MARTIN, E., DENISSOVA, L., KOZLOV, M., MARKOVTSOV, V., KASHLEV, M., HEUMANN, H., NIKIFOROV, V., GOLDFARB, A. & MUSTAEV, A. 1996. Mapping of Catalytic Residues in the RNA Polymerase Active Center. *Science,* 273**,** 107-109.

ZENKIN, N., KULBACHINSKIY, A., BASS, I. & NIKIFOROV, V. 2005b. Different rifampin sensitivities of Escherichia coli and Mycobacterium tuberculosis RNA polymerases are not explained by the difference in the beta-subunit rifampin regions I and II. *Antimicrobial agents and chemotherapy,* 49**,** 1587-1590.

ZENKIN, N., KULBACHINSKIY, A., YUZENKOVA, Y., MUSTAEV, A., BASS, I., SEVERINOV, K. & BRODOLIN, K. 2007. Region 1.2 of the RNA polymerase sigma subunit controls recognition of the -10 promoter element. *Embo j,* 26**,** 955-64.

ZENKIN, N., NARYSHKINA, T., KUZNEDELOV, K. & SEVERINOV, K. 2006. The mechanism of DNA replication primer synthesis by RNA polymerase. *Nature,* 439**,** 617-620.

ZENKIN, N. & SEVERINOV, K. 2004. The role of RNA polymerase σ subunit in promoter-independent initiation of transcription. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 4396-4400.

ZHANG, G., CAMPBELL, E. A., MINAKHIN, L., RICHTER, C., SEVERINOV, K. & DARST, S. A. 1999. Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution. *Cell,* 98**,** 811-824.

ZHANG, Y., DEGEN, D., HO, M. X., SINEVA, E., EBRIGHT, K. Y., EBRIGHT, Y. W., MEKLER, V., VAHEDIAN-MOVAHED, H., FENG, Y., YIN, R., TUSKE, S., IRSCHIK, H., JANSEN, R., MAFFIOLI, S., DONADIO, S., ARNOLD, E. & EBRIGHT, R. H. 2014a. GE23077 binds to the RNA polymerase 'i' and 'i+1' sites and prevents the binding of initiating nucleotides. *eLife,* 3**,** e02450.

ZHANG, Y., FENG, Y., CHATTERJEE, S., TUSKE, S., HO, M. X., ARNOLD, E. & EBRIGHT, R. H. 2012. Structural Basis of Transcription Initiation. *Science,* 338**,** 1076-1080.

ZHANG, Z., DAI, F., LUO, F., ZHONG, M., HUANG, Z., HOU, T. & XU, J. 2014b. Could high-concentration rifampicin kill rifampicin-resistant M. tuberculosis? Rifampicin MIC test in rifampicin-resistant isolates from patients with osteoarticular tuberculosis. *Journal of Orthopaedic Surgery and Research,* 9**,** 124.

ZHU, W., HAUPENTHAL, J., GROH, M., FOUNTAIN, M. & HARTMANN, R. W. 2014. New insights into the bacterial RNA polymerase inhibitor CBR703 as a starting point for optimization as an anti-infective agent. *Antimicrob Agents Chemother,* 58**,** 4242-5.

ZUMLA, A., GEORGE, A., SHARMA, V., HERBERT, R. H. N., OXLEY, A. & OLIVER, M. 2015. The WHO 2014 Global tuberculosis report—further to go. *The Lancet Global Health,* 3**,** e10-e12.