TRANSCRIPTION
2021, VOL. 12, NO. 4, 92–102
https://doi.org/10.1080/21541264.2021.1964328

Taylor & Francis
Taylor & Francis Group

REVIEW

RNA polymerases from low G+C gram-positive bacteria

Michael Miller [a], Aaron J. Oakley[a], and Peter J. Lewis [a,b]

[a]School Of Environmental And Life Sciences, University Of Newcastle, Callaghan, NSW, Australia; [b]School Of Chemistry And Molecular Bioscience, University Of Wollongong And Illawarra Health And Medical Research Institute, Wollongong, Nsw, Australia

ABSTRACT

The low G + C Gram-positive bacteria represent some of the most medically and industrially important microorganisms. They are relied on for the production of food and dietary supplements, enzymes and antibiotics, as well as being responsible for the majority of nosocomial infections and serving as a reservoir for antibiotic resistance. Control of gene expression in this group is more highly studied than in any bacteria other than the Gram-negative modelÂ Escherichia coli, yet until recently no structural information on RNA polymerase (RNAP) from this group was available. This review will summarize recent reports on the high-resolution structure of RNAP from the model low G + C representativeÂ Bacillus subtilis, including the role of auxiliary subunits **δ** and **ε**, and outline approaches for the development of antimicrobials to target RNAP from this group.

ARTICLE HISTORY
Received 3 June 2021
Revised 28 July 2021
Accepted 31 July 2021

KEYWORDS
RNA polymerase; Transcription regulation; Low G+C bacteria; Auxiliary factors; Antibiotics

Introduction

RNAP structures from multiple bacterial species have been determined. Within molecular microbiology, two organisms have dominated research into transcription and its regulation. *Escherichia coli*, a Gram-negative bacterium, has been extensively studied to understand fundamental mechanistic aspects of transcription and its regulation *e.g.* [1]. *Bacillus subtilis*, a Gram-positive organism, has been extensively studied for the regulatory processes associated with the initiation of differential gene expression followed by compartment-specific transcription activation and gene expression during the developmental process of sporulation [2].

At a structural level, RNAP from *E. coli* ($RNAP_{Eco}$) has been studied for many years and with the advent of modern cryo electron microscopy and single particle analysis techniques (cryoEM) [3], was one of the first pseudo-atomic resolution multi-subunit RNAP structures solved [4]. Prior to the advent of these current resources, high resolution structural data from X-ray crystallography was largely obtained from thermophiles for which little molecular biology data on transcription and its regulation had been performed. Knowing the structure of $RNAP_{Eco}$ allowed the reconciliation of structural and functional data in one system which has enabled profound new insights into the mechanisms of transcription and its regulation *e.g* [5–8].

Such structural data on the Gram-positive *B. subtilis* system has been lagging, but high resolution structures of several important complexes of *B. subtilis* RNAP ($RNAP_{Bsu}$) have been recently published that enable a similar reconciliation of structural and molecular data [9–11]. Despite the considerable similarity between all multi-subunit RNAPs from bacteria, there are important mechanistic differences that can now be examined. For example, initiation complexes tend to undergo multiple rounds of abortive initiation prior to leaving the promoter region and entering the elongation phase in *E. coli*, but similar effects are not observed in *B. subtilis* [12,13]. The concentration and identity of the initiating NTP ([iNTP]) also has a major effect on transcription efficiency in *B. subtilis* [14].

Increasing our understanding of transcription regulation through structure-function studies is particularly important as *B. subtilis* is an industrially significant organism used in the production of enzymes (proteases, lipases, amylases), surfactants, and antibiotics (bacitracin) which

CONTACT Peter J. Lewis peter.lewis@newcastle.edu.au School Of Environmental And Life Sciences, University Of Newcastle, Callaghan, Nsw, 2308, Australia

have highly complex regulatory circuits controlling the expression of their genes. As a member of the *Firmicutes*, it is closely related to many of the most important clinical pathogens such as *Staphylococcus, Streptococcus, Enterococcus*, and *Clostridia*. In this review we will examine the structure of $RNAP_{Bsu}$ and compare it with that of other bacteria, focussing on its unique features and auxiliary subunits and include reference to homologous RNAPs from closely related *Firmicutes* pathogens and how this information could be exploited in structure-based drug design.

Overall structure: RNA polymerases from the low G + C firmicutes

All bacterial RNAPs have a similar overall subunit composition comprising two α subunits that form an asymmetric dimer scaffold upon which the catalytic β and β' subunits assemble (Figure 1). Whilst there is high sequence conservation of residues/motifs involved in core polymerase function there are many sites where lineage-specific inserts are found which substantially alter the size and shape of RNAP from different organisms [15]. $RNAP_{Bsu}$ has two lineage-specific inserts (βIn5 and β696–705) in the β subunit and often, but not always, an additional small ε subunit (Figure 1(a)). Sequence analysis indicates the 86 amino acid βIn5 insertion is restricted to the phylum *Firmicutes*. The location of inserts in the β and β' subunits of *Thermus thermophilus* ($RNAP_{Tth}$), *Mycobacterium smegmatis* ($RNAP_{Msm}$) and *E. coli* ($RNAP_{Eco}$) relative to the structure of $RNAP_{Bsu}$ are shown in red in (Figure 1(b-d)), respectively. Due to the lack of lineage-specific inserts (see below) $RNAP_{Bsu}$ is more compact (shorter and narrower) than other bacterial RNAPs: 150 Å × 112 Å × 123 Å

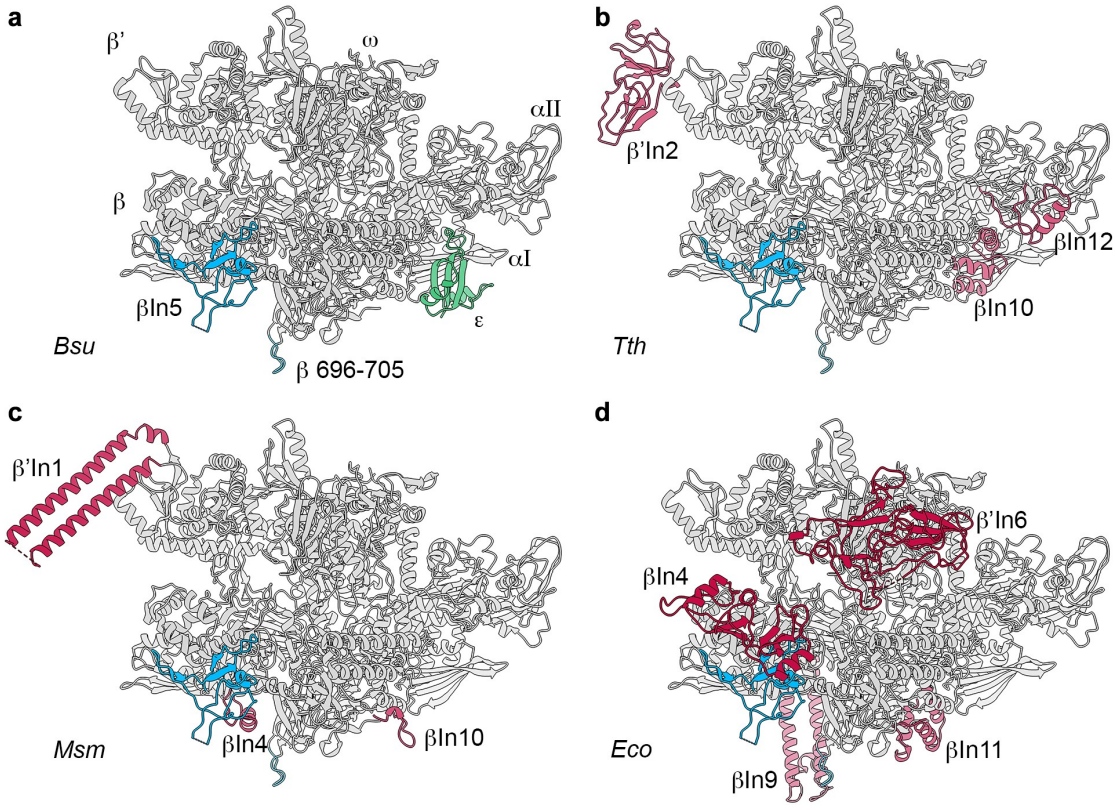

Figure 1. Comparison of bacterial RNAPs. panel A shows the structure of *B. subtilis* RNAP elongation complex (*Bsu* PDB ID 6WVJ). nucleic acids have been removed for clarity. *B. subtilis* lineage specific inserts βIn5 and β696–705 are shown in blue, and the ε subunit in green. since the ε subunit is absent in other organisms and is not always present in all RNAP complexes isolated from *B. subtilis* it is not shown in subsequent panels. In panels B-D $RNAP_{Bsu}$ is shown with lineage-specific inserts from superimposed RNAPs from *thermus thermophilus* (*tth* PDB ID 2O5I), *Mycobacterium smegmatis* (*Msm* PDB ID 6F6W) and *E. coli* (*Eco* PDB ID 6ALF), respectively, labeled and shown in red.

(L × W × H), *vs* 157 × 153 × 136 Å; $RNAP_{Eco}$, 183 × 107 × 115 Å; *Mycobacterium smegmatis* ($RNAP_{Msm}$), and 170.1 × 110.1 × 127.8 Å; *Thermus thermophilus* ($RNAP_{Tth}$) [10].

Transcription initiation involves the RNAP holoenzyme, promoter DNA, and transcription factors (TFs). The primary housekeeping σ transcription factor in *Bacillus subtilis*, σ^A, contains sub-domains $\sigma_{1.1}$, σ_2, $\sigma_{3.1}$, $\sigma_{3.2}$, and σ_4. The holoenzyme structure (core + σ^A) shown in (Figure 2(a)) is based on the complex (PDB ID 7CKQ) with multidrug resistance regulator BmrR [9] and shows an open complex (DNA strands separated at the −10 region, red oval, and template strand inserted in the active site within the primary channel). In this structure much of the lineage-specific βln5 insertion (see below) is missing, but the flexible β-flap tip, important in binding essential transcription factor NusA [16], is visible and binds to the σ_4 domain that also binds the −35 promoter sequence (red circle, Figure 2(a)). In combination with the transcription elongation complex (EC) [10], a near complete structure of the *B. subtilis* enzyme can be modeled. The location of the primary (DNA binding) and secondary (NTP entry) channels are marked as dashed circles in the EC and shown in (Figure 2(b)). It should be noted that the α-C terminal domain that interacts with both transcription factors and DNA sequences was absent in all $RNAP_{Bsu}$ structures reported to date, due to the highly flexible sequence connecting the N- and C-terminal domains, and lack of DNA/transcription factor for the α-C terminal domain to bind to. The structure of this important DNA

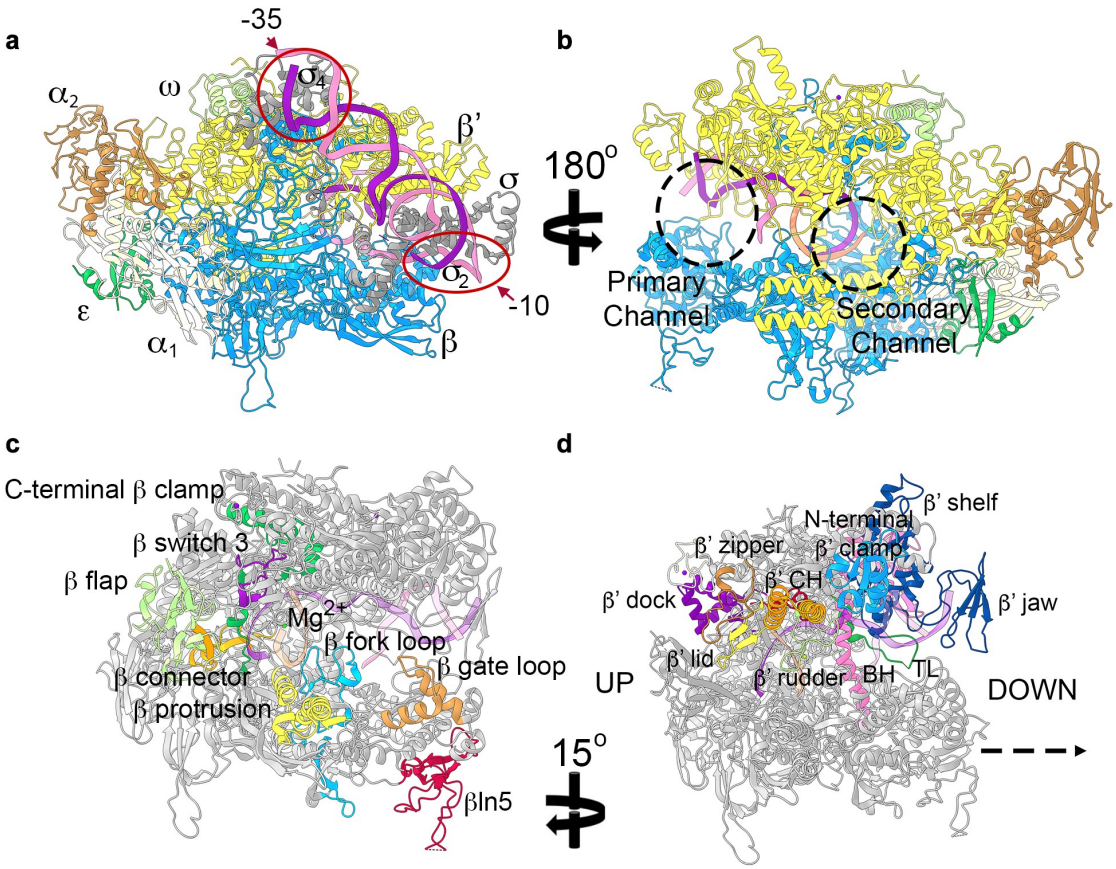

Figure 2. Structure of *B. sutbtilis* RNAP. panel A shows the structure of RNAP holoenzyme (PDB ID 7CKQ), and panel B an elongation complex (PDB ID 6WVJ). subunit coloring; α_1 cream, α_2 brown, β blue, β′ yellow, ε dark green, ω pale green, σ gray. Template strand DNA is shown in dark purple, non-template strand DNA in pink, and RNA in orange. The −10 and −35 promoter elements are ringed in panel A, and the primary and secondary channels circled in panel B. panels C and D show key functional elements of the β and β′ subunits, respectively, as defined by [11]. β subunit elements in panel B; C-terminal β clamp dark green, β switch 3 purple, β flap pale green, β connector vermillion, β protrusion yellow, β fork loop blue, β gate loop orange, βln5 red. β′ subunit elements in panel C; β′ dock purple, β′ lid yellow, β′ zipper orange and clamp helix (CH) orange, β′ rudder pale green, β′ clamp blue, β′ shelf and jaw dark blue, bridge helix (BH) pink, trigger loop (TL) dark green. Up, and downstream sides of RNAP are indicated for reference.

and protein interaction domain has been solved in complex with the transcription factor Spx [17] enabling inclusion in models based on transcription initiation or antitermination complex structures from other organisms [18,19] making it possible to further augment the structures presented here.

A small, highly conserved, ω subunit binds around the C-terminus of the β' subunit and is associated with enhancing subunit folding and incorporation of the β' subunit into the core structure [20,21]. Database searches using the Conserved Domain Architecture Retrieval Tool (CDART) identified 911 sequences from the *Bacilli* and *Lactobacilli*, but not *Clostridia*, corresponding to an additional small subunit named ε, previously annotated as a second ω subunit (dark green, Figure 1a, Figure 2a and Figure 2b), who's function has proven to be enigmatic [22–24]. Determination of the X-ray crystal structure of ε revealed it showed remarkable similarity to the *E. coli* phage T7 gp2 that binds the β' jaw region of RNAP (labeled in Figure 2(d)) where it inhibits transcription initiation by the host cell RNAP [22,25]. Based on this similarity, it was hypothesized that ε could be involved in phage protection through binding to the β' jaw of $RNAP_{Bsu}$ [22,24]. Determination of the structure of *B. subtilis* holoenzyme and a transcription recycling complex comprising core RNAP with the ATP-dependent remodeling factor HelD revealed ε bound in a pocket formed mainly by the β' and α_1 subunits behind the secondary channel on the downstream side of the enzyme (Figures 1(a), Figures 2b [9–11]). The location of ε overlaps that of the βln10 and −12 inserts of $RNAP_{Tth}$ (Figure 1(b)) as well as a region of Archaeal/Eukaryotic Pol II Rpo3/RPB3 associated with enzyme stability, suggesting a similar role for ε in organisms, such as *B. subtilis*, that are meso-thermophilic and capable of vigorous growth up to ~53°C [10].

Major functional motifs in the β and β' subunits common to other bacterial RNAPs that are important in DNA strand separation and rewinding (β fork loop,β'-rudder, -lid, -zipper) DNA clamping (β-protrusion, -gate loop, β' clamp helices, N-terminal β' clamp), DNA binding cleft flexibility (β switch 3) *etc.* are all highly conserved in the *B. subtilis* enzyme and are labeled in Figure 2(c-d)) [10,11]. However, given the differences in aspects of transcription such as open complex stability (above) between *B. subtilis* and *E. coli*, sequence differences and/or insertions (*e.g.* the β'ln6 insertion into the trigger-loop in *E. coli*, Figure 1(d)), these regions remain important areas of focus in structure-function studies in the *Firmicutes*.

Across the eubacteria the β and β' subunits contain lineage-specific inserts that are largely of unknown function [15]. Despite the role of most of these lineage-specific inserts being unknown, the βln9 insert in the β subunit of $RNAP_{Eco}$ has recently been putatively implicated in coupling transcription and translation under certain conditions [26]. Elucidation of the structure of $RNAP_{Bsu}$ revealed the structure of the only major lineage-specific insert in *Firmicutes* RNAPs (labeled βln5 in Figure 2(c) [10]. Reconciliation of data from previous studies indicates that the βln5 insert is involved in binding to the C-terminal tudor domain of helicase/translocase PcrA [27,28] that is known to interact strongly with RNAP [27,29–31], and is important in the resolution of R-loops (DNA-RNA hybrids) that form as a result of high transcriptional activity [32]. The βln5 insert is located within the major lobe of the β subunit which is one of the least conserved regions in bacterial RNAPs and is the site of many other lineage-specific inserts [15] raising the possibility this part of RNAP may be important in providing a platform for lineage-specific transcription factor interaction modules. Overall, due to the lack of lineage-specific inserts (excepting βln5), RNAP from *B. subtilis* and other low G + C Gram-positive bacteria represent the smallest multi-subunit RNAPs: *B. subtilis* $\alpha_2\beta\beta'\omega\varepsilon$, 352.32 kDa; *Mycobacterium tuberculosis* $\alpha_2\beta\beta'\omega$, 363.19 kDa; *cf. E. coli* $\alpha_2\beta\beta'\omega$, 389.05 kDa.

The δ subunit

Many *Firmicutes* also contain a small δ subunit that is tightly associated with and present at approximately equimolar concentrations with respect to RNAP [29,33]. Detailed analysis, to be published elsewhere, indicates that δ is found in nearly all of the *Firmicutes* apart from the *Clostridia*, as well as *Mycoplasma*. δ is a bipartite protein of 173 amino acids with a globular

N-terminal domain and unstructured highly acidic C-terminal domain of approximately equal sizes [34,35]. It has been implicated in multiple regulatory roles associated with transcription initiation, inhibition of nonspecific transcription, transcription complex recycling and transcription termination [24,36–41]. The binding site of δ on RNAP has been the subject of considerable speculation. Independent studies placed it adjacent to the RNA exit channel [42] or close to/inside the DNA binding cleft [43]. Determination of the structure of a *B. subtilis* RNAP-HelD-δ recycling complex by Pei *et al.* [11], showed δ binds on the β' subunit close to the DNA binding cleft, consistent with findings of *in vivo* cross-linking mass spectrometry studies [43], reconciling the observed biochemical effects of δ on transcription with the structure of a δ-containing transcription complex. In subcellular localization studies of δ and RNAP in dual fluorescent protein labeled cells, δ perfectly colocalised with RNAP and was present at similar levels suggesting it is associated with RNAP throughout all stages (initiation, elongation, termination) of the transcription cycle [33]. Assuming the N-terminal domain of δ remains bound around the β' jaw/N-terminal β' clamp region and the acidic unstructured C-terminal domain is both mobile and flexible, we may propose a mechanism for modulation of transcription (Figure 3). A closed initiation complex, based on the $RNAP_{Bsu}$ open complex structure [9] in which the N-terminal $\sigma_{1.1}$ domain [44] that competes with nucleic acid binding in the primary channel can be modeled *in situ* based on equivalent structures from *E. coli* [25] is shown in Figure 3(a)). As the transcription initiation complex transitions from a closed to open complex, the $\sigma_{1.1}$ domain swings out of the primary channel as unwound DNA enters, placing the single stranded template strand with the transcription start site nucleotide in position for base pairing with the initiating nucleotide triphosphate [9,45].

Upon $\sigma_{1.1}$ dissociation during the formation of the open complex, the δ C-terminal domain would be able to access DNA within the primary channel (Figure 3(b)). The highly negative charge of the δ C-terminal domain would encourage dissociation of weakly bound DNA (poor/nonspecific promoter sequences bound by σ-factors) [24,36]. To examine the role of δ in termination, the cryo EM structure of the $RNAP_{Eco}$ paused EC (PDB ID 6FLP) was used as a template to model

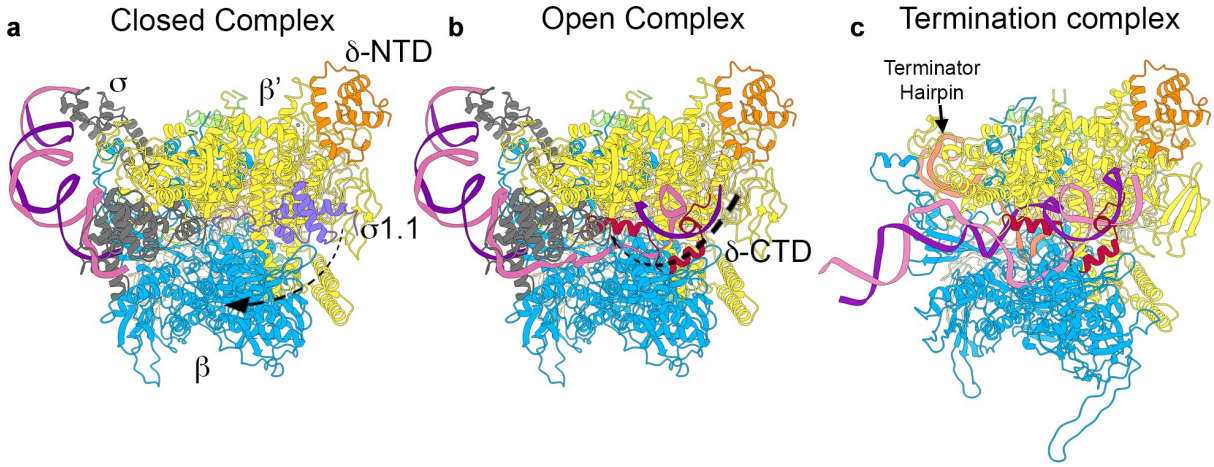

Figure 3. A model for δ subunit activity during transcription initiation and termination. In all panels, RNAP subunits are colored as in Figure 1 with the addition of σ region 1.1 (σ1.1) shown in lavender, the N-terminal domain of δ (δ-NTD) in orange, and the C-terminal domain (δ-CTD) in red. panel A shows a closed transcription initiation complex with the δ-NTD bound around the β' shelf. The CTD is not shown (see text for details). σ1.1 is shown in primary channel with the arrow indicating the dissociation of σ1.1 from this site during the transition from a closed to open initiation complex. panel B shows an open complex in which promoter DNA strands have separated and the template strand has moved into the active site within the primary channel. δ-CTD is shown moving into the primary channel where, due to its polyanionic nature it can compete with DNA in the primary channel helping to prevent transcription initiation from cryptic/weak promoters. panel C shows a model of a transcription termination complex with a RNA hairpin (terminator hairpin). in the RNA exit channel. The polyanionic δ-CTD is able to disrupt the RNA-DNA hybrid upstream from the active site aiding dissociation of RNA from the complex and RNAP recycling following termination of transcription.

termination hairpin RNA in the $RNAP_{Bsu}$ EC. The model suggests that the highly flexible C-terminal domain of δ can interact with both DNA and the RNA transcript in the active site (Figure 3(b-c)), consistent with the observation that δ is much more efficient at displacing RNA from an EC than it is DNA [37]. In pause/termination complexes, the negatively charged δ C-terminal domain would aid the dissociation of RNA, facilitating transcription termination and transcription complex recycling (Figure 3(c)) [11,36,41].

Comparison of $RNAP_{Bsu}$ with RNAP from other firmicutes

B. subtilis itself is a biologically and industrially significant organism, being important in soil health and promotion of plant growth, protection from plant pathogens, the industrial production of enzymes (*e.g.* proteases/amylases), supplements (*e.g.* nicotinic acid), and antibiotics (*e.g.* bacitracin), as a probiotic, as a foodstuff (*e.g.* in natto), and in the study of regulation of gene expression, especially during cellular differentiation in sporulation [2]. As the most studied member of the low G + C Gram-positive *Firmicutes* it is also an important model, and closely related to major pathogens including *B. cereus, B. anthracis, Staphylococcus sp., Streptococcus sp., Enterococcus sp*. and *Clostridium sp*. Organisms such as the *Enterococci* are associated with dissemination of antibiotic resistance determinants, and many clinical isolates of *Staphylococcus* now carry resistance to one or more antibiotics. *S. aureus* is commensal in about 30% of the population and an opportunistic pathogen that remains a significant burden on health systems through nosocomial infections, which has been exacerbated in recent years by the rise of community-acquired infections (especially methicillin-resistant; MRSA infections) [46]. Organisms such as *C. difficile* are associated with diseases difficult to treat successfully with many antibiotics (*e.g. C. difficile* associated diarrhea (CDAD) relapse is common following vancomycin treatment), and represent a significant burden in terms of both morbidity and mortality to health systems [47]. While fidaxomicin was approved by the FDA for treatment of CDAD in 2011, resistance to this drug (lipiarmycin) was first reported in 1977 [48,49], and it is clear new derivatives are needed, as well as a new arsenal of novel compounds to slow the rise of antibiotic resistant infections.

Sequence alignment of RNAP subunits from representatives of these organisms was performed using *B. subtilis* 168, *S. aureus* USA300, *E. faecalis* V583, *S. pyogenes* M1 GAS, *C. difficile* 630 and *C. perfringens* 13 sequences and the resulting CLUSTAL alignment outputs mapped onto the *B. subtilis* elongation complex (PDB ID 6 WVJ) in ChimeraX [50,51] with the nucleic acids removed for clarity. The resulting homology sequence maps are shown in Figure 4 along with a phylogenetic tree produced in MrBayes [52,53] for the *rpoC* (β' subunit) using the *E. coli rpoC* sequence as an outlier to root the tree. The bootstrap probability values of 1 indicate absolute confidence in the branch divisions and lengths, and agree perfectly, as expected, with the segregation of *B. subtilis* and *S. aureus* to the *Bacilli, S. pyogenes* and *E. faecalis* to the *Lactobacilli*, and *C. difficile* and *C. perfringens* to the *Clostridia*.

The level of sequence conservation is high, especially in the β and β' clamps, active site and β flap where the majority of the functional motifs (see Figure 2(c-d)) are found. Although the βln5 is present in all of the organisms from which sequences were selected, the level of sequence conservation is relatively low in the major β lobe and βln5 region (Figure 4(c)), consistent with the hypothesis (above) that this region maybe important for providing class/order/species-specific binding sites for transcription factors. *B. subtilis*-specific sequences that are absent in the other organisms, such as the 10 amino acid insert at β E696-G705 that protrudes from the bottom of the structure are shown in yellow (Figure 3(b-c)). Given the industrial and medical importance of this group of bacteria, regions of identity/difference can be targeted in functional studies or exploited in the rational design of inhibitor compounds as new antimicrobial leads.

Antimicrobial development options

Transcription is an underutilized target for new antibiotic development, although significant efforts are currently underway to identify promising new leads and to improve the properties of

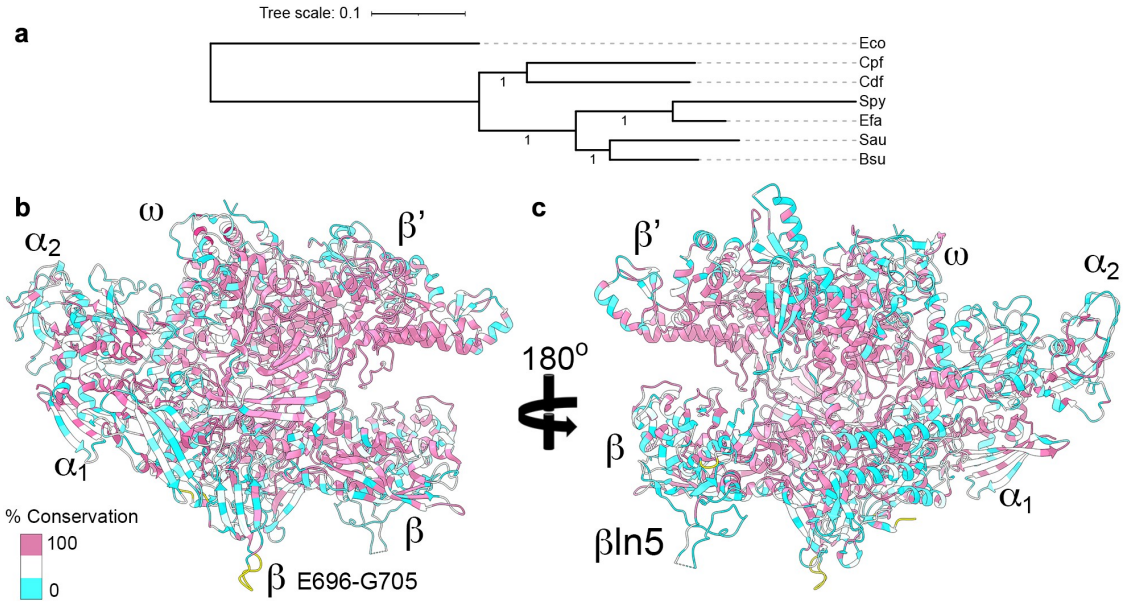

Figure 4. Mapping sequence conservation of pathogenic *Firmicutes* to *B. subtilis* RNAP. Panel A shows a bayesian tree of sequence alignments of the β subunit of *B. subtilis* (Bsu), *S. aureus* (Sau), *E. faecalis* (Efa), *S. pyogenes* (Spy), *C. difficile* (Cdf), and *C. perfringens* (Cpf). The sequence of *E. coli* β subunit (Eco) was used to root the tree. tree scale represents amino acid substitutions per site. bootstrap values are shown on the branches. Panels B and C show up- and downstream views of RNAP, respectively. subunits are labeled as well as the common to *Firmicutes* βIn5 insert, and *B. subtilis*-specific β E696-G705 insert (yellow). A color scale for sequence conservation shown on the structures is shown on the bottom left of Panel B with sequences 100% conserved pink, 0% conserved cyan and >0, <100 in white.

existing clinical compounds [48,54]. Many promising compounds fail to make it to market as broad spectrum antibiotics due to the problems of identifying hits that are able to cross the outer membrane of Gram-negative bacteria, despite showing excellent activity against Gram-positives. Infections due to the *Firmicutes*, including *S. aureus, C. difficile*, vancomycin-resistant *Enterococcus*, and drug-resistant *Streptococcus* (Group A and B), have been identified by the Center for Disease Control (CDC) as organisms of major clinical concern for which new approaches/treatments for infection are required [47], and there is a case for developing more narrow spectrum drugs that target this group [55].

Fidaxomicin is a semi-synthetic macrolide that inhibits transcription initiation and was approved for use in CDAD in 2012. *C. difficile* (and other *Clostridia*) are exquisitely sensitive to fidaxomicin, but this is not a property shared by most other *Firmicutes* with some *Streptococci* being >500 × more resistant than *C. difficile* [56]. Structures of fidaxomicin in complex with RNAP from *Mycobacterium tuberculosis* have been solved [57,58] enabling modeling of the drug bound to the *B. subtilis* holoenzyme (PDB ID 7CKQ) with sequence alignments to pathogenic *Firmicutes* homology mapped as in the previous section (Figure 5(a)). Other than the homodichloro-orsellenyl moiety that is adjacent to the β-flap tip and −35 promoter sequence binding σ region 4, the bulk of fidaxomicin is buried within the enzyme at the base of the RNAP clamp. All of the RNAP and σ sequences the remaining bulk of fidaxomicin interacts with (switch regions Sw2, Sw3, Sw4 and the σ finger [57,58]); are highly conserved (pink, Figure 5(a), right side box) consistent with the broad spectrum activity of this compound in *in vitro* transcription assays (although *E. coli* holoenzyme is quite resistant to fidaxomicin [57]). Increasing the spectrum of activity of fidaxomicin may depend on improving cell permeability properties, especially for organisms such as *S. pneumoniae*, where production of a capsule layer may inhibit efficient cell penetration.

An area of antimicrobial research that is showing promise for compounds highly active against *Firmicutes* is the development of compounds that inhibit the essential interaction between RNAP and σ^A. Establishment of a functional holoenzyme complex is dependent on the interaction between a small

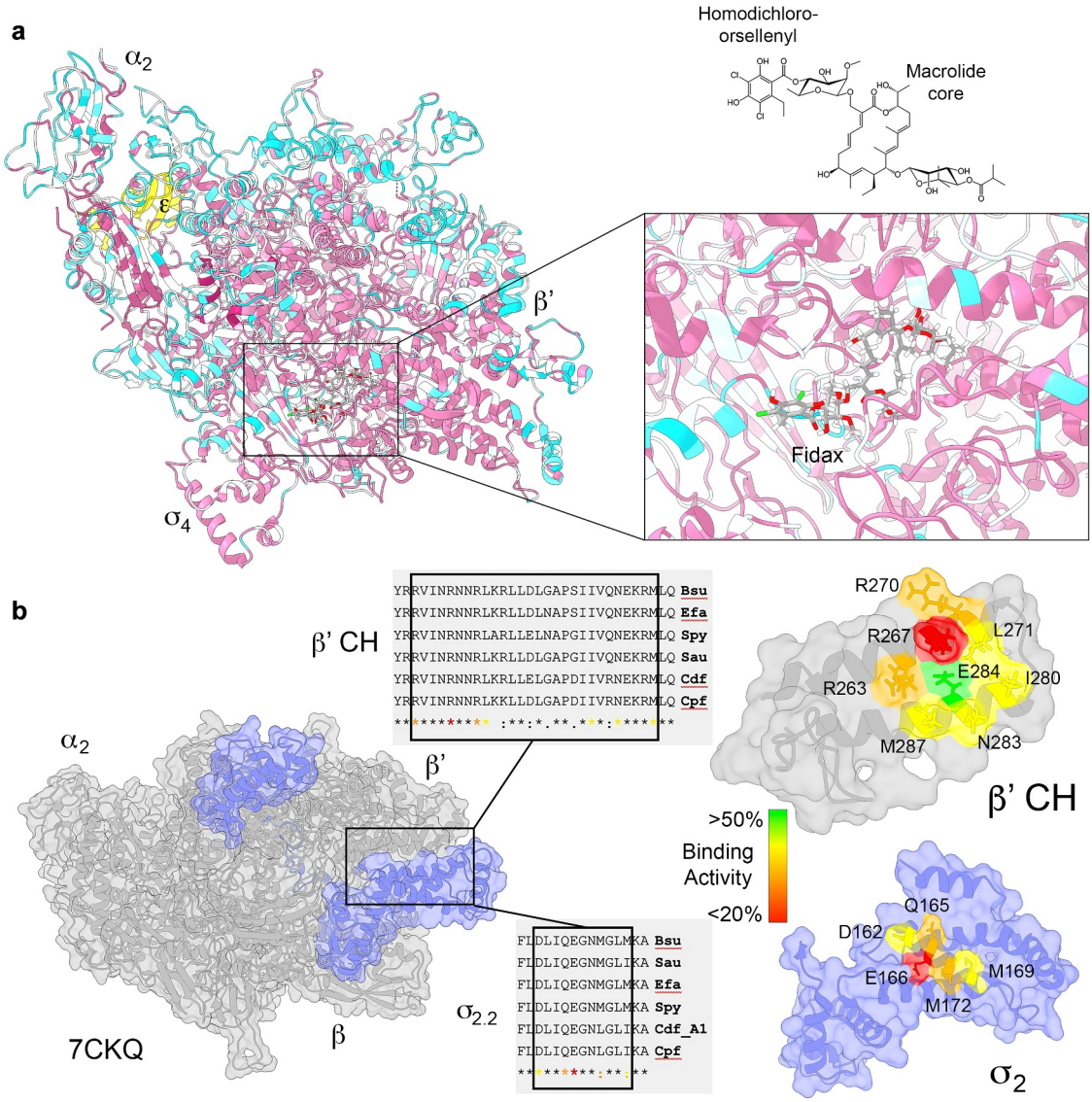

Figure 5. Transcription inhibition drug targets in *firmicutes* RNAP. panel A shows sequence conservation mapping of pathogenic *Firmicutes* with coloring as in Figure 4 mapped onto the *B. subtilis* RNAP holoenzyme (PDB ID 7CKQ). The α_2, and β′ subunits and σ_4 domain are labeled for reference. The ε subunit which is not conserved in *Clostridia* is shown in yellow. Fidaxomicin (Fidax) is shown docked in the holoenzyme structure (box), which is shown in an enlarged box on the right (see text for details). The structure of fidaxomicin is shown above the right hand box approximately aligned for reference with the homodichloro-orsellenyl and macrolide core labeled for reference. Panel B, left, shows *B. subtilis* holoenzyme with core subunits colored gray and σ^A in pale blue. The box indicates the interaction site between the β′ CH and $\sigma_{2.2}$ regions essential for formation of holoenzyme. sequence alignments of the relevant regions are shown adjacent to the holoenzyme. Note, *C. difficile* encodes two σ^A subunits, but only the alignment for *sigA1* is shown as *sigA2* is not expressed to a significant level during vegetative growth. Strain labeling is the same as in Figure 4. The right hand side shows enlarged regions of the β′ CH and $\sigma_{2.2}$ regions with amino acids involved in formation of the holoenzyme color coded according to their importance as determined from mutagenesis studies [59]. The color ramp indicates the relative binding activity mutation causes to holoenzyme formation.

region of the $\sigma_{2.2}$ region with the β clamp helix (CH) [59,60]. These sequences are highly conserved in the *Firmicutes* (Figure 4(b)), and across the eubacteria, making this an excellent target for the development of molecules that inhibit this essential protein–protein interaction (PPI) [48,61]. Multiple research programs involving high-throughput small molecule screens, structure-based drug design, and small peptide antagonists have yielded promising results [48,62–70]. PPIs are an attractive target for drug development as simultaneous complementary mutations are required at two unlinked genetic loci to confer

resistance to a compound whilst retaining the interaction, which has the potential to substantially reduce the rate at which resistance develops [71].

Mutagenesis studies to quantify the importance of specific amino acid residues in formation of holoenzyme have been determined (Figure 5(b) right [59]), that have enabled the construction of pharmacophores for screening compound libraries for potential inhibitor molecules. A major issue with such a target is that the interaction site between $\sigma_{2.2}$ and the β' clamp helix is relatively flat making binding specificity and avidity potentially problematic. Nevertheless, compounds have been developed that are highly specific for bacterial initiation complexes, showing no binding activity against human RNAP, and that target initiation complex formation in live bacterial cells as determined by cytological assay [48]. Whilst many of these compounds show limited or no activity against Gram-negative bacteria, excellent results have been obtained against Gram-positive bacteria, including those carrying resistance to multiple antibiotics providing an avenue for development of new drugs effective against Gram-positive pathogens [48,65,68,69].

Concluding statement

Determination of the structure of RNAP from *B. subtilis* now opens the way to undertake detailed structure-function studies on the mechanism of transcription in this Gram-positive model, particularly with respect to mechanistic aspects that are different to those in *E. coli*. The presence of the δ and ε subunits in these structures will also enable structure-function studies on the roles of these *Firmicutes*-specific proteins, and the lack of lineage-specific inserts provides a "minimal" RNAP structure for functional studies on the inserts found in RNAPs from other organisms. Finally, the development of antimicrobial compounds that specifically target transcription is an expanding field, and we can now undertake structure-guided development of new lead antibiotics to combat infections caused by the *Firmicutes*.

Acknowledgments

This work was supported by funding from the Australian Research Council (DP210100365) to PJL and AJO. MM was supported by a PhD scholarship from the University of Newcastle Priority Research Centre for Drug Discovery.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Australian Research Council [DP210100365].

ORCID

Michael Miller http://orcid.org/0000-0001-7990-3381
Peter J. Lewis http://orcid.org/0000-0002-1992-062X

References

[1] Browning DF, Busby SJ. Local and global regulation of transcription initiation in bacteria. Nat Rev Microbiol. 2016;14(10):638–650.
[2] Errington J, Aart LTV. Microbe Profile: bacillus subtilis: model organism for cellular development, and industrial workhorse. Microbiology (Reading). 2020;166(5):425–427.
[3] Benjin X, Ling L. Developments, applications, and prospects of cryo-electron microscopy. Protein Sci. 2020;29(4):872–882.
[4] Chen J, Wassarman KM, Feng S, et al. 6S RNA Mimics B-Form DNA to Regulate Escherichia coli RNA Polymerase. Mol Cell. 2017;68(2):e386.
[5] Guo X, Myasnikov AG, Chen J, et al. Structural Basis for NusA Stabilized Transcriptional Pausing. Mol Cell. 2018;69(816–827):e814.
[6] Abdelkareem M, Saint-Andre C, Takacs M, et al. Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. Mol Cell. 2019;75(2):e294.
[7] Kang JY, Llewellyn E, Chen J, et al. Structural basis for transcription complex disruption by the Mfd translocase. Elife. 2021;10.
[8] Said N, Hilal T, Sunday ND, *et al.* Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase rho. Science. 2021;371.
[9] Fang C, Li L, Zhao Y, *et al.* The bacterial multidrug resistance regulator BmrR distorts promoter DNA to activate transcription. Nat Commun. 2020;11(1):6284.
[10] Newing TP, Oakley AJ, Miller M, et al. Molecular basis for RNA polymerase-dependent transcription complex

recycling by the helicase-like motor protein HelD. Nat Commun. 2020;11(1):6420.
[11] Pei HH, Hilal T, Chen ZA, *et al*. The delta subunit and NTPase HelD institute a two-pronged mechanism for RNA polymerase recycling. Nat Commun. 2020;11 (1):6418.
[12] Ishikawa S, Oshima T, Kurokawa K, et al. RNA polymerase trafficking in bacillus subtilis cells. J Bacteriol. 2010;192(21):5778–5787.
[13] Mooney RA, Davis SE, Peters JM, et al. Regulator trafficking on bacterial transcription units in vivo. Mol Cell. 2009;33(1):97–108.
[14] Krasny L, Gourse RL. An alternative strategy for bacterial ribosome synthesis: bacillus subtilis rRNA transcription regulation. EMBO J. 2004;23(22):4473–4483.
[15] Lane WJ, Darst SA. Molecular evolution of multisubunit RNA polymerases: structural analysis. J Mol Biol. 2010;395(4):686–704.
[16] Ma C, Mobli M, Yang X, et al. RNA polymerase-induced remodelling of NusA produces a pause enhancement complex. Nucleic Acids Res. 2015;43(5):2829–2840.
[17] Newberry KJ, Nakano S, Zuber P, et al. Crystal structure of the Bacillus subtilis anti-alpha, global transcriptional regulator, Spx, in complex with the alpha C-terminal domain of RNA polymerase. Proc Natl Acad Sci U S A. 2005;102(44):15839–15844.
[18] Hudson BP, Quispe J, Lara-Gonzalez S, et al. Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. Proc Natl Acad Sci U S A. 2009;106(47):19830–19835.
[19] Krupp F, Said N, Huang YH, et al. Structural basis for the action of an all-purpose transcription anti-termination factor. Mol Cell. 2019;74(1):e145.
[20] Ghosh P, Ishihama A, Chatterji D. Escherichia coliRNA polymerase subunit ω and its N-terminal domain bind full-length β' to facilitate incorporation into the α2 β subassembly. Eur J Biochem. 2001;268(17):4621–4627.
[21] Minakhin L, Bhagat S, Brunning A, et al. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A. 2001;98(3):892–897.
[22] Keller AN, Yang X, Wiedermannova J, et al. epsilon, a new subunit of RNA polymerase found in gram-positive bacteria. J Bacteriol. 2014;196(20):3622–3632.
[23] Spiegelman GB, Hiatt WR, Whiteley HR. Role of the 21,000 molecular weight polypeptide of Bacillus subtilis RNA polymerase in RNA synthesis. J Biol Chem. 1978;253(6):1756–1765.
[24] Weiss A, Shaw LN. Small things considered: the small accessory subunits of RNA polymerase in Gram-positive bacteria. FEMS Microbiol Rev. 2015;39 (4):541–554.
[25] Bae B, Davis E, Brown D, et al. Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of sigma70 domain 1.1. Proc Natl Acad Sci U S A. 2013;110(49):19772–19777.
[26] Wang C, Molodtsov V, Firlar E, et al. Structural basis of transcription-translation coupling. Science. 2020;369 (6509):1359–1365.
[27] Harriott K (2012), The University of Newcaslte.
[28] Sanders K, Lin CL, Smith AJ, et al. The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase. Nucleic Acids Res. 2017;45 (7):3875–3887.
[29] Delumeau O, Lecointe F, Muntel J, et al. The dynamic protein partnership of RNA polymerase in Bacillus subtilis. Proteomics. 2011;11(15):2992–3001.
[30] Gwynn EJ, Smith AJ, Guy CP, et al. The conserved C-terminus of the PcrA/UvrD helicase interacts directly with RNA polymerase. PLoS One. 2013;8(10):e78141.
[31] Noirot-Gros MF, Dervyn E, Wu LJ, et al. An expanded view of bacterial DNA replication. Proc Natl Acad Sci U S A. 2002;99(12):8342–8347.
[32] Urrutia-Irazabal I, Ault JR, Sobott F, et al. Analysis of the PcrA-RNA polymerase complex reveals a helicase interaction motif and a role for PcrA/UvrD helicase in the suppression of R-loops. Elife. 2021;10. DOI:10.7554/eLife.68829
[33] Doherty GP, Fogg MJ, Wilkinson AJ, et al. Small subunits of RNA polymerase: localization, levels and implications for core enzyme composition. Microbiology. 2010;156(12):3532–3543.
[34] Kuban V, Srb P, Stegnerova H, *et al*. Quantitative Conformational Analysis of Functionally Important Electrostatic Interactions in the Intrinsically Disordered Region of Delta Subunit of Bacterial RNA Polymerase. J Am Chem Soc. 2019;141:16817–16828.
[35] Motackova V, Sanderova H, Zidek L, et al. Solution structure of the N-terminal domain of Bacillus subtilis δ subunit of RNA polymerase and its classification based on structural homologs. Proteins. 2010;78 (7):1807–1810.
[36] Juang YL, Helmann JD. The delta subunit of Bacillus subtilis RNA polymerase. An allosteric effector of the initiation and core-recycling phases of transcription. J Mol Biol. 1994;239(1):1–14.
[37] Lopez de Saro FJ, Woody AY, Helmann JD. Structural analysis of the Bacillus subtilis delta factor: a protein polyanion which displaces RNA from RNA polymerase. J Mol Biol. 1995;252(2):189–202.
[38] Lopez de Saro FJ, Yoshikawa N, Helmann JD. Expression, abundance, and RNA polymerase binding properties of the delta factor of Bacillus subtilis. J Biol Chem. 1999;274(22):15953–15958.
[39] Nicolas P, Mader U, Dervyn E, *et al*. Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. Science. 2012;335 (6072):1103–1106.
[40] Rabatinova A, Sanderova H, Jirat Matejckova J, et al. The delta subunit of RNA polymerase is required for

rapid changes in gene expression and competitive fitness of the cell. J Bacteriol. 2013;195 (11):2603–2611.
[41] Wiedermannova J, Sudzinova P, Koval T, *et al.* Characterization of HelD, an interacting partner of RNA polymerase from Bacillus subtilis. Nucleic Acids Res. 2014;42(8):5151–5163.
[42] Prajapati RK, Sengupta S, Rudra P, et al. Bacillus subtilis delta factor functions as a transcriptional regulator by facilitating the open complex formation. J Biol Chem. 2016;291(3):1064–1075.
[43] de Jong L, de Koning EA, Roseboom W, *et al.* In-Culture Cross-Linking of Bacterial Cells Reveals Large-Scale Dynamic Protein-Protein Interactions at the Peptide Level. J Proteome Res. 2017;16 (7):2457–2471.
[44] Zachrdla M, Padrta P, Rabatinova A, et al. Solution structure of domain 1.1 of the sigma(A) factor from Bacillus subtilis is preformed for binding to the RNA polymerase core. J Biol Chem. 2017;292 (28):11610–11617.
[45] Murakami KS, Darst SA. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol. 2003;13 (1):31–39.
[46] Tong SY, Davis JS, Eichenberger E, et al. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev. 2015;28:603–661.
[47] CDC. (2013), *Centre for Disease Control.*
[48] Ma C, Yang X, Lewis PJ. Bacterial Transcription as a Target for Antibacterial Drug Development. Microbiol Mol Biol Rev. 2016;80(1):139–160
[49] Sonenshein AL, Alexander HB, Rothstein DM, et al. Lipiarmycin-resistant ribonucleic .acid polymerase mutants of Bacillus subtilis. J Bacteriol. 1977;132(1):73–79.
[50] Goddard TD, Huang CC, Meng EC, et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. 2018;27(1):14–25.
[51] Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: structure Visualization for Researchers, Educators, and Developers. Protein Sci. 2020.
[52] Huelsenbeck JP, Ronquist F. MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics. 2001;17(8):754–755.
[53] Ronquist F, Teslenko M, Van Der Mark P, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–542.
[54] Maffioli SI, Zhang Y, Degen D, *et al.* Antibacterial Nucleoside-Analog Inhibitor of Bacterial RNA Polymerase. Cell. 2017;169(7):e1223.
[55] Alm RA, Lahiri SD. Narrow-Spectrum Antibacterial Agents-Benefits and Challenges. Antibiotics (Basel). 2020;9.
[56] Goldstein EJ, Babakhani F, Citron DM. Antimicrobial activities of fidaxomicin. Clin Infect Dis. 2012;55(Suppl suppl_2):S143–148.
[57] Boyaci H, Chen J, Lilic M, et al. Fidaxomicin jams Mycobacterium tuberculosis RNA polymerase motions needed for initiation via RbpA contacts. Elife. 2018;7.
[58] Lin W, Das K, Degen D, *et al.* Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). Mol Cell. 2018;70(1):60–71.e15.
[59] Johnston EB, Lewis PJ, Griffith R. The interaction of Bacillus subtilis σ A with RNA polymerase. Protein Sci. 2009;18(11):2287–2297.
[60] Arthur TM, Anthony LC, Burgess RR. Mutational analysis of beta '260–309, a sigma 70 binding site located on Escherichia coli core RNA polymerase. J Biol Chem. 2000;275(30):23113–23119.
[61] Cossar PJ, Lewis PJ, McCluskey A. Protein-protein interactions as antibiotic targets: a medicinal chemistry perspective. Med Res Rev. 2020;40(2):469–494.
[62] Andre E, Bastide L, Michaux-Charachon S, et al. Novel synthetic molecules targeting the bacterial RNA polymerase assembly. J Antimicrob Chemother. 2006;57 (2):245–251.
[63] Glaser BT, Bergendahl V, Thompson NE, et al. LRET-based HTS of a small-compound library for inhibitors of bacterial RNA polymerase. Assay Drug Dev Technol. 2007;5(6):759–768.
[64] Husecken K, Negri M, Fruth M, et al. Peptide-Based Investigation of the Escherichia coli RNA Polymerase σ 70 :Core Interface As Target Site. ACS Chem Biol. 2013;8(4):758–766.
[65] Ma C, Yang X, Kandemir H, et al. Inhibitors of bacterial transcription initiation complex formation. ACS Chem Biol. 2013;8(9):1972–1980.
[66] Mielczarek M, Thomas RV, Ma C, et al. Synthesis and biological activity of novel mono-indole and mono-benzofuran inhibitors of bacterial transcription initiation complex formation. Bioorg Med Chem. 2015;23(8):1763–1775.
[67] Wenholz DS, Zeng M, Ma C, et al. Small molecule inhibitors of bacterial transcription complex formation. Bioorg Med Chem Lett. 2017;27 (18):4302–4308.
[68] Ye J, Chu AJ, Harper R, *et al.* Discovery of Antibacterials That Inhibit Bacterial RNA Polymerase Interactions with Sigma Factors. J Med Chem. 2020;63 (14):7695–7720.
[69] Ye J, Chu AJ, Lin L, et al. Benzyl and benzoyl benzoic acid inhibitors of bacterial RNA polymerase-sigma factor interaction. Eur J Med Chem. 2020;208:112671.
[70] Sartini S, Levati E, Maccesi M, *et al.* New Antimicrobials Targeting Bacterial RNA Polymerase Holoenzyme Assembly Identified with an in Vivo BRET-Based Discovery Platform. ACS Chem Biol. 2019;14:1727–1736.
[71] Igler C, Rolff J, Regoes R. Multi-step vs. single-step resistance evolution under different drugs, pharmacokinetics and treatment regimens. Elife. 2021;10. DOI:10.7554/eLife.64116