Slowing down to get ahead: functional and structural characterisation of toxin-antitoxin systems from Mycobacterium tuberculosis

USHER, BEN

Slowing down to get ahead: functional and structural characterisation of toxin-antitoxin systems from *Mycobacterium tuberculosis*

Ben Usher

BSc (Hons), MScR

University College

Thesis for the degree of Doctor of Philosophy in Biological Sciences, Department of Biosciences, Durham University, 2021

Abstract

Toxin-antitoxin (TA) systems are ubiquitously found encoded on bacterial chromosomes and mobile genetic elements. They comprise a toxin protein, which interferes with an essential cellular process to inhibit growth, and an antitoxin, either sRNA or protein, which neutralises toxicity. TA modules are implicated in various roles, including plasmid maintenance, phage defence, and aiding pathogenicity. *Mycobacterium tuberculosis* boasts the greatest number of TA systems, with at least eighty identified thus far. The causative agent of tuberculosis, *M. tuberculosis* infects one third of the global population and was responsible for 1.4 million deaths in 2019 alone. Elucidating TA mechanisms that contribute to *M. tuberculosis* pathogenicity may help inform strategies to control future infections.

This study has characterised a family of four TA systems identified in *M. tuberculosis*: $MenA_1$-$MenT_1$, $MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$. The MenT toxins are members of the widespread nucleotidyltransferase-like DUF1814 protein family. Functional tests showed that $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ inhibit growth in *E. coli* through a reversible mechanism. The X-ray crystallographic structures of the $MenT_1$, $MenT_3$, and $MenT_4$ toxins were solved to 1.65 Å, 1.78 Å, and 1.23 Å, respectively. An additional $MenT_3$ structure was solved to 1.59 Å which was phosphorylated at S78, an important residue for $MenA_3$ antitoxicity. $MenT_1$, $MenT_3$ and $MenT_4$ are bi-lobed globular proteins which feature a shared toxin fold and conserved active site. The crystal form of complexed $MenA_1$:$MenT_1$ was also solved to 1.44 Å, which shows $MenA_1$ binding asymmetrically across two $MenT_1$ protomers. Protein interaction studies indicated that the MenA-MenT family is comprised of multiple TA classes. Finally, biochemical assays demonstrated that $MenT_3$ and $MenT_4$ inhibit protein synthesis *in vitro*, implicating translation as the cellular target. This study's characterisation of the MenA-MenT family expands our knowledge of *M. tuberculosis* TA systems, and helps reveal a novel mechanism by which a toxin inhibits bacterial growth.

Declaration

This thesis is the outcome of my own work and is submitted solely for the degree of Doctor of Philosophy in Biological Sciences. Any results that arose from collaborative efforts are specifically mentioned in-text.

All research presented herein, with the exception of collaborative contributions where stated, were performed in the laboratory of Dr Tim R Blower, Department of Biosciences, Durham University, during the period August 2017 to June 2021.

Ben Usher

December 2021

Statement of Copyright

Contents

List of Figures

List of Tables

Key of Abbreviations

AA	amino acid(s)
AaRS	aminoacyl tRNA synthetase
Abi	abortive infection
ADP	adenosine diphosphate
AEC	anion exchange chromatography
AMR	antimicrobial resistant/resistance
Ap	ampicillin
AT	antitoxin
ATP	adenosine triphosphate
β-gal	β-galactosidase
bp	base pair(s)
BSA	buried surface area
CFU	colony forming unit(s)
Cm	chloramphenicol
COG	cluster of orthologous group
CTD	C-terminal domain
CTP	cytidine triphosphate
Da	dalton
D-glu	D-glucose
dH_2O	distilled water
DHFR	dihydrofolate reductase
DMSO	dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
DUF	domain of unknown function
EDTA	ethylenediaminetetraacetic acid
EOP	efficiency of plating
ES+ TOF MS	positive electrospray ionisation time-of-flight mass spectrometry
EtOH	ethanol
Fwd	forward
GNAT	GCN5-related *N*-acetyltransferase
GTP	guanosine triphosphate

h	hour(s)
hSENP2	human sentrin/SUMO-specific protease 2
ICE	integrative and conjugative element
IDP	intrinsically disordered protein
IDR	intrinsically disordered region
IPTG	isopropyl β-D-thiogalactopyranoside
IR	inverted repeat
kb	kilobase(s)
kDa	kilodalton(s)
Km	kanamycin
L-ara	L-arabinose
LB	Luria broth
LBA	Luria broth agar
LIC	ligation independent cloning
M9A	M9 minimal agar
M9M	M9 minimal media
MAD	multi-wavelength anomalous diffraction
mAU	milli-arbitrary unit(s)
MDR	multi-drug resistant/resistance
MDR-TB	multi-drug resistant tuberculosis
MDT	multi-drug tolerant/tolerance
MenA	mycobacterial AbiE-like NTase antitoxin
MenT	mycobacterial AbiE-like NTase toxin
MES	2-(N-morpholino) ethanesulfonic acid
MGE	mobile genetic element
MIC	minimum inhibitory concentration
min	minute(s)
MQ	Milli-Q (Merck)-purified and deionised water
MR	molecular replacement
mRNase	mRNA-specific ribonuclease
MS	mass spectrometry
MTBC	*M*. *tuberculosis* complex
NTase	nucleotidyltransferase
NTD	N-terminal domain
NTP	nucleoside triphosphate

OD_{420}	optical density at wavelength 420 nm
OD_{600}	optical density at wavelength 600 nm
ONPG	o-nitrophenyl-β-D-galactopyranoside
ORF	open reading frame
P	promoter
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIPES	1,4-piperazinediethanesulfonic acid
PMF	proton-motive force
PNK	polynucleotide kinase
(p)ppGpp	guanosine tetraphosphate/pentaphosphate
PSK	post-segregational killing
PTM	post-translational modification
$_R$	resistance/resistant
RBS	ribosome binding site
Rev	reverse
RMSD	root mean square deviation
RNase	ribonuclease
ROS	reactive oxygen species
RR-TB	rifampicin-resistant tuberculosis
σ	sigma
$_S$	sensitivity/sensitive
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulphate
sec	second(s)
SEC	size exclusion chromatography
SeMet	selenomethionine
SEP	phosphoserine
Ser	serine
SHX	serine hydroxymate
SI	super-integron
SOB	super optimal broth
Sp	spectinomycin

SR	stringent response
SRL	Sarcin-Ricin loop
sRNA	small RNA
ssDNA	single-stranded DNA
T	toxin
TA	toxin-antitoxin
TAE	tris-acetate-EDTA
TB	tuberculosis
Tc	tetracycline
TCS(s)	two-component system(s)
TEMED	*N*,*N*,*N*',*N*'-tetramethylethylenediamine
Thr	threonine
TraSH	transposon site hybridisation
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UTP	uridine triphosphate
UTR	untranslated region
v/v	volume/volume
WHO	World Health Organisation
wHTH	winged helix-turn-helix
WT	Wild-Type
w/v	weight/volume
XDR-TB	extensively drug-resistant tuberculosis

Acknowledgements

First and foremost, I thank my supervisor Tim for his outstanding support throughout this project. I have been fortunate to be one of his early PhD students (trial runs), and have benefited enormously from his guidance, enthusiasm, and *nearly* limitless patience. It has been a genuine pleasure working under his supervision.

I also thank the funding bodies that facilitated this project – The Academy of Medical Sciences and Durham University – and once again gratefully acknowledge Tim's role in securing this funding.

I wish to further thank the PIs at Durham for their various words of wisdom and advice (Dr Gary Sharples and Prof. Ehmke Pohl), our collaborators further afield in France (Prof. Pierre Genevaux and Dr Yiming Cai) and New Zealand (Prof. Peter Fineran) for their respective roles, and my colleagues in the Blower Lab for helping make the office and lab an entertaining place to work.

Thanks go to various members of the non-academic staff of the Chemistry Department, for providing a pleasant and friendly workplace, and offering life-saving cheap (and eventually free) tea/coffee.

Finally, I thank my parents for their constant support, and Alex, for her enduring patience and for being there.

Publications

The results of this thesis have so far directly contributed to two peer-reviewed articles. A third peer-reviewed article (in *Journal of Structural Biology*) was published during the course of this project as the result of data from a side-project, that will not be discussed in this thesis. The publications are listed below, and are laid out in full in the Appendix of this thesis:

- Izaak N. Beck, Ben Usher, Hannah G. Hampton, Peter C. Fineran, Tim R. Blower (2020) Antitoxin autoregulation of *M. tuberculosis* toxin-antitoxin expression through negative cooperativity arising from multiple inverted repeat sequences. *Biochem. J.* **477**(12): 2401–2419

- Yiming Cai*, Ben Usher*, Claude Gutierrez, Anastasia Tolcan, Moise Mansour, Peter C. Fineran, Ciarán Condon, Olivier Neyrolles, Pierre Genevaux†‡ and Tim R. Blower†‡ (2020) A nucleotidyltransferase toxin inhibits growth of Mycobacterium tuberculosis through inactivation of tRNA acceptor stems. *Sci. Adv.* **6**(31): eabb6651

- Ben Usher, Nils Birkholz, Izaak N. Beck, Robert D. Fagerlund, Simon A. Jackson, Peter C. Fineran, Tim R. Blower (2021) Crystal structure of the anti-CRISPR repressor Aca2. *J. Struct. Biol.* **213**(3): 107752

1. Introduction

1.1. Toxin-antitoxin systems

Toxin-antitoxin (TA) systems are typically small, bicistronic loci found ubiquitously in bacteria, archaea, and some fungi, encoded on chromosomes and mobile genetic elements (MGEs)[1–9]. Their distribution amongst different organisms varies; *Escherichia coli* (*E. coli*), the best-studied organism in relation to TA systems, contains at least 39[10], whilst *Mycobacterium tuberculosis* (*M. tuberculosis*) contains as many as 88[11,12]. In contrast, non-pathogenic soil-dwelling *Mycobacterium smegmatis* (*M. smegmatis*) harbours just five TA modules[13]. TA loci are generally (though not universally) characterised by an antitoxin gene encoded immediately upstream of a cognate toxin gene, both of which are co-transcribed from a single promoter[14–16]. This is reflected in the common nomenclature for individual TA systems, where the antitoxin typically precedes the toxin, a convention which will be used herein for consistency.

Expression of the toxin produces a protein capable of inhibiting growth of the host cell by interfering with an essential cellular process, for example translation or DNA replication[7,17]. Antitoxins, in contrast, are more varied, and manifest either as proteins or small, non-coding RNAs (sRNA)[7]. Under normal growth conditions, the cognate antitoxin acts antagonistically to either directly or indirectly neutralise toxin activity, facilitating unhindered bacterial growth[7,17]. The mechanisms underlying toxin activation remain a point of debate: it has previously been suggested that under certain stress conditions, the more labile antitoxin is preferentially degraded by cellular proteases, which frees and activates the toxin to inhibit growth[18]. This is proposed to reduce metabolic burden and induce a stress-tolerant state which promotes survival until favourable conditions return[18,19]. However, this model has been questioned in recent years, with new reports suggesting that stress-induced antitoxin degradation is not the sole driver of toxin activation, and that stress does not guarantee toxin-mediated growth arrest[20,21].

TA systems were first identified in *E. coli*, where the CcdA-CcdB (control of cell death) system was observed to function in plasmid maintenance through a process termed post-segregational killing (PSK)[1,22–24] (Figure 1.1). The presence of the *ccdA-ccdB* system on the *E. coli* F plasmid ensures that following cell division, only plasmid-containing progeny harbouring the full TA locus can survive. In plasmid-deficient daughter cells, the cellular pool of CcdA antitoxin is readily degraded by Lon protease[24,25], and with no means to replenish antitoxin levels, the CcdB toxin is free to induce cell death by poisoning DNA gyrase[1,22,26–29].

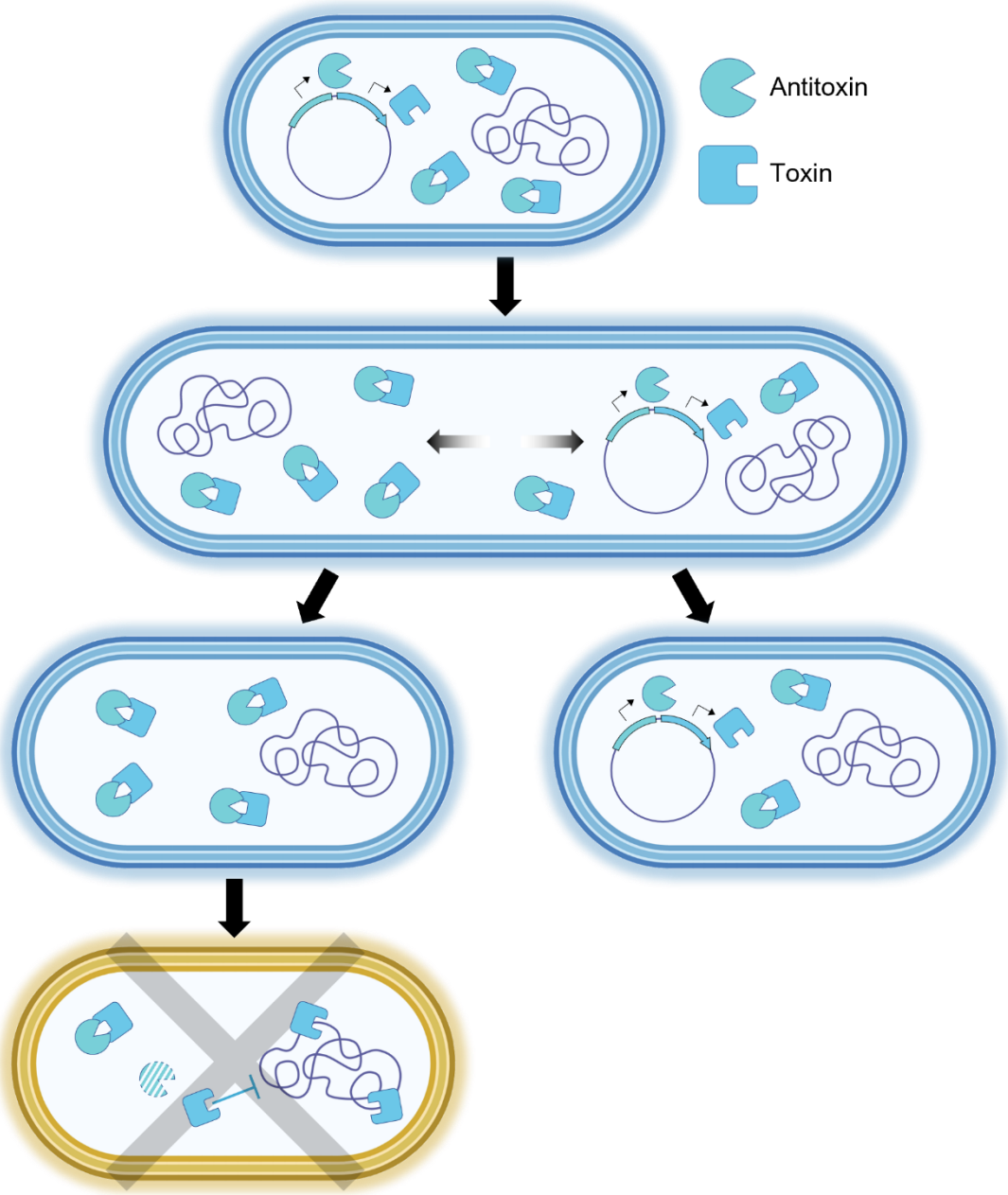


Figure 1.1. Schematic of TA-mediated post-segregational killing (PSK). During normal cell growth, the presence of the TA-encoding plasmid facilitates regeneration of the antitoxin, thereby providing continuous neutralisation of the cognate toxin. Following cell division, progeny lacking the TA plasmid, or with an impaired antitoxin gene, are unable to synthesise new antitoxin. Due to its relative lability, the antitoxin is quickly degraded by cellular proteases, freeing the toxin to inhibit growth and induce cell death.

Numerous TA systems have since been found conserved in a broad range of bacteria, and implicated in a variety of roles. For instance, whilst plasmid-borne TA systems might typically mediate PSK[1,2,30], chromosomal TA systems are often associated with cell survival and the bacterial response to stress, for example protecting against nutrient starvation and predation by bacteriophages[16,31,32]. They have also been linked to other roles which are more hotly debated, for example bacterial persistence[7,33,34]. To date, seven different types of TA system have so far been classified, based primarily on the biological function of the antitoxin and the mechanism of toxin neutralisation.

1.1.1. Type I TA systems

The antitoxins of the type I TA systems are characteristically antisense sRNAs that inhibit toxin activity by base-pairing with toxin mRNA, preventing its translation[9] (Figure 1.2A). The toxin, on the other hand, is typically a small, membrane lytic peptide which depolarises the bacterial membrane to inhibit ATP synthesis[9]. The first identified member of this TA family was the Sok-Hok system from the *parB* locus of plasmid R1, involved in plasmid maintenance via PSK[2,30,35]. This TA locus encodes the Sok (suppression of killing) antitoxin, an unstable antisense sRNA, the Hok (host killing) toxin, a short hydrophobic peptide, and the *mok* (mediation of killing) reading frame, which overlaps and is translationally coupled to *hok*, and features complementarity to Sok-sRNA[30,36,37]. The *hok* mRNA exists in two forms: the primary form; a full length, translationally inactive mRNA – which conceals the complementary *mok* sequence and therefore binds Sok-sRNA poorly – is activated by slow 3' processing to generate the second form; a translationally active truncated mRNA which is recognised and bound by Sok[38–41]. The primary transcript form accumulates in cells due to its inert nature and inability to be inactivated by Sok[35]. In *sok-hok* plasmid-carrying cells, processed and truncated mature *hok* mRNA is bound by Sok-sRNA; the resulting RNA duplex interferes with ribosome binding and leads to rapid cleavage by RNase III[37,39,42,43]. In plasmid-free cells, unstable Sok-sRNA is degraded, and with no means to replenish it, the available reservoir of *hok* mRNA is freely processed and translated, with the toxin at liberty to target the bacterial membrane[36].

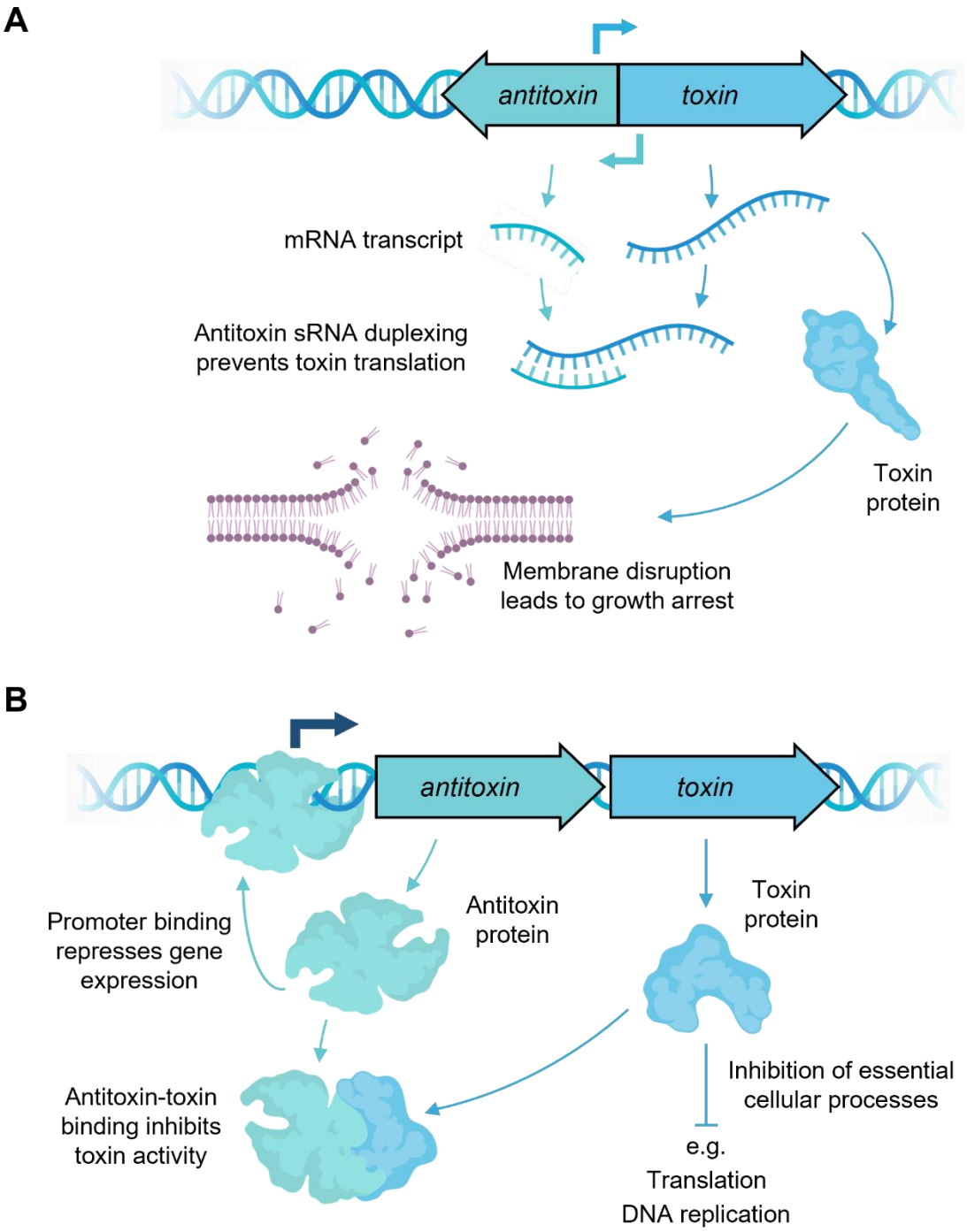


Figure 1.2. Schematic models of canonical Type I and II TA systems. (A) The type I TA system. The antitoxin comprises an antisense sRNA which base-pairs with a short region of the full-length mRNA transcript, preventing its translation. When processed, the hydrophobic toxin protein functions to disrupt membranes, causing growth arrest. **(B)** The type II TA system. The antitoxin protein typically regulates activity via a two-pronged approach: a DNA-binding domain facilitates promoter binding and transcriptional regulation, whilst a second domain is able to directly bind the toxin, interfering with its activity. The toxin protein inhibits growth by targeting essential processes such as DNA replication and translation, and in many cases acts as a co-repressor of transcription alongside the cognate antitoxin.

Following the identification of the Sok-Hok system, many homologous systems and additional type I families have been discovered spread throughout Enterobacteriaceae, Vibrionaceae and Firmicutes lineages[44–48], with type I loci often found repeated within the same genome[49]. The first type I TA system found in Gram-positive bacteria was RNAII-Fst, located in the plasmid partitioning (*par*) locus on plasmid pAD1 from *Enterococcus faecalis*[50,51], with many related TA loci identified in other Gram-positive bacteria[52]. Other examples of well-characterised and extensively found type I families include RatA-TxpA, first reported in *Bacillus subtilis*[53], and the RdlD-LdrD, OhsC-ShoB, IstR-TisAB and SibC-IbsC systems, originally discovered in *E. coli*[54–56]. In addition, type I TA systems are not limited to membrane disruption as a mode of toxicity: the SOS-induced SymE toxin has homology to the type II antitoxin MazE, but functions like a type II toxin as a ribonuclease (RNase), whilst the RalR toxin is a non-specific DNase; each is inhibited by the cognate antitoxin sRNA, SymR or RalA respectively, as per a canonical type I system[57,58].

Interestingly, the evolution and distribution of type I TA systems is unlike that of type II, with limited evidence for horizontal gene transfer, but rather for lineage-specific duplication events[45]. These duplications are evolutionarily stable, suggesting a functional role for these TA loci in their respective organisms[45]. However, the relatively narrow distribution of known type I modules may in part be a relic of the difficulty in predicting and detecting these small genetic elements using *in silico* means[59]. Questions also surround the evolution of the antitoxin sRNAs based on the divergence of type I genetic architectures: typical type I sRNAs are encoded opposite either the toxin gene's 5'-untranslated region (UTR) or 3'-UTR, whereas other type I families feature the antitoxin encoded divergently to the toxin, but still maintaining regions of complementarity[45,49,60].

1.1.2. Type II TA systems

Type II TA systems are the largest and most widely studied class of TA system, consisting of a two-gene operon encoding a labile antitoxin protein and a relatively stable toxin protein (Figure 1.2B). Type II antitoxins typically contain two functional domains which contribute a dual-faceted approach to regulating toxicity. One is a protein-binding domain which specifically binds the cognate toxin to inhibit its activity and negate toxicity[61] (Figure 1.2B). Antitoxicity can arise from occlusion of the toxin active sites, as with the PezA-PezT system[62], or via allosteric regulation, such as with the HipB-HipA system, where HipB sequesters HipA and locks it into an inactive conformation[63]. The second domain is a DNA-binding domain

which can engage specific elements within its own promoter to autoregulate TA transcription[61] (Figure 1.2B). In many cases the toxins also act as co-repressors, physically interacting with the antitoxins to enhance DNA binding and tighten transcription repression[61] (See, *1.5*).

The first type II systems were discovered on plasmids, involved in maintaining plasmid stability via PSK; indeed, the CcdA-CcdB system, described above, is a type II TA module whose toxin component poisons DNA gyrase to induce DNA cleavage[64]. Many type II systems have since been linked to a range of physiological roles such as stress response, phage inhibition, biofilm formation, and persistence[61,65]. Although several type II toxins have been described that target DNA replication, such as CcdB and ParE, type II toxins predominantly target protein synthesis, either as general RNases degrading mRNA transcripts (mRNases), or via inhibition of translation machinery[61,64,66,67].

Type II toxin families are usually defined by structural similarity, with at least nine super-families currently recognised: AtaT/TacT, MazF, ParE/RelE, HicA, HipA, VapC, FicT/Doc, MbcT and Zeta[61]. They are broadly distributed, with multiple homologues often existing in the same chromosome; *M. tuberculosis*, for example, harbours approximately 50 VapB-VapC family members alone[11]. Interestingly, the sequence similarity of toxins within these superfamilies is often poor. This may account for the notably "relaxed" mechanistic commonality among toxins of the same family: for example, the FicT family member Doc is a kinase which inhibits translation, whereas other FicT toxins disrupt DNA topology by adenylating DNA gyrase and topoisomerase IV[68–70]. Similarly, RelE mRNase toxins from distantly related species have evolved different cleavage specificities, whereas the MazF super-family members MazF and CcdB are mRNases and gyrase poisons, respectively[64,71,72].

1.1.3. Type III TA systems

The first identified and best characterised type III TA system was ToxI-ToxN, encoded on the pECA1039 cryptic plasmid from *Pectobacterium atrosepticum* (*P. atrosepticum*), a Gram-negative phytopathogen[31]. Two 2009 studies demonstrated that ToxI-ToxN also functions as an Abi (abortive infection) system, protecting against bacteriophages via a reversible bacteriostatic mechanism[31,73]. The *toxN* gene encodes a protein toxin able to inhibit bacterial growth, whilst the *toxI* gene comprises a sequence of 5.5 almost identical repeats of 36 nucleotides which encode the sRNA antitoxin[31,73]. The *toxI-toxN* locus is bicistronic, co-

transcribed from a single promoter, and has a Rho-independent transcriptional terminator situated between the two genes leading to excess ToxI-sRNA relative to ToxN[31] (Figure 1.3A).

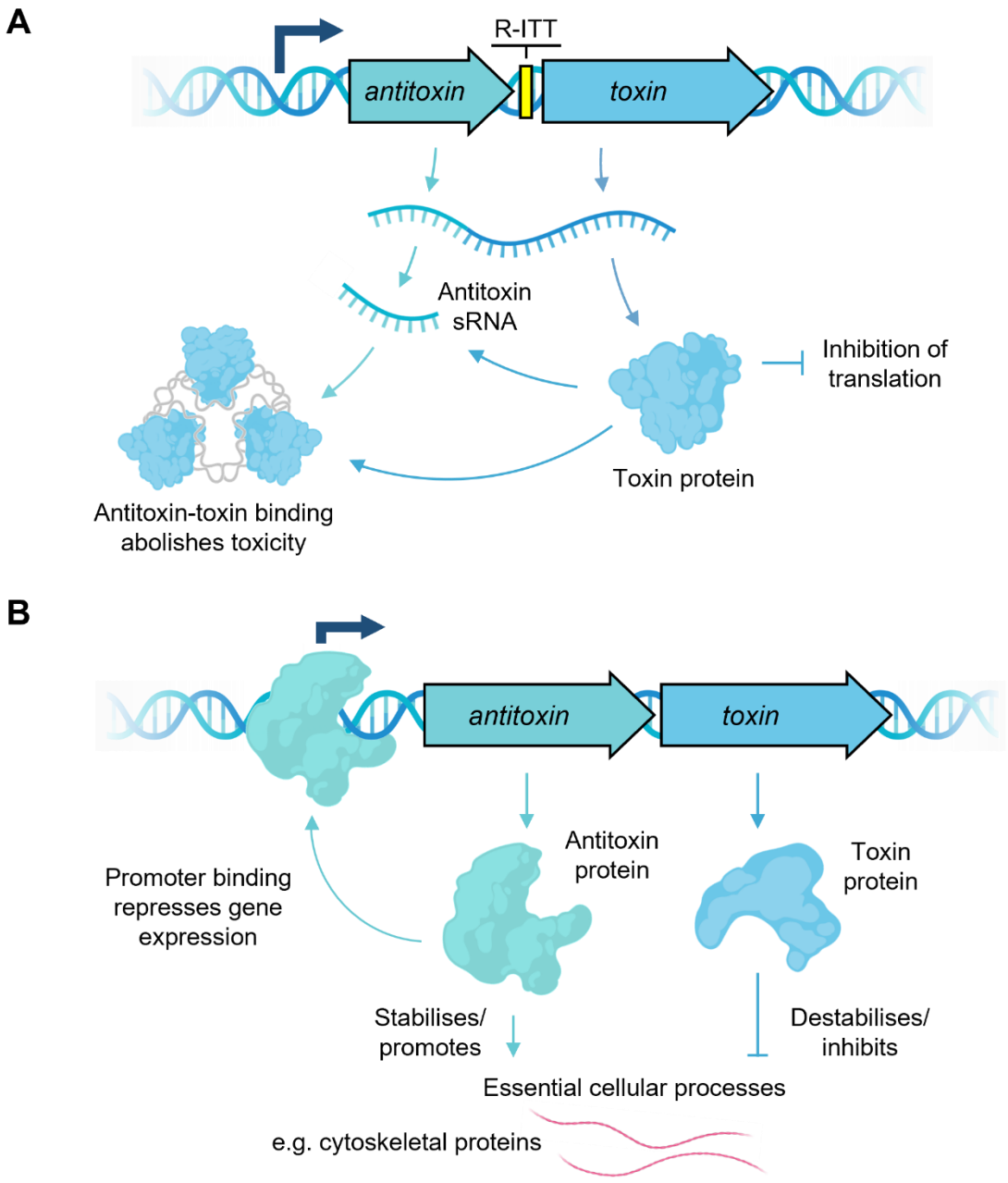


Figure 1.3. Schematic models of canonical Type III and IV TA systems. (A) The type III TA system. The toxin protein is an RNase which interferes with translation by cleaving mRNA transcripts. The antitoxin is processed by the toxin RNase to produce sRNA. This forms pseudoknots which bind the toxin directly, inhibiting its activity. A Rho-independent transcriptional terminator (R-ITT) sits between the two genes, resulting in greater antitoxin expression relative to the toxin. **(B)** The type IV TA system. Antitoxin and toxin proteins do not directly interact, but instead target the same cellular process; where the toxin stalls growth, for example by destabilising bacterial filaments, the antitoxin negates toxicity by antagonistically stabilising these proteins. Further TA regulation can occur at the transcriptional level, mediated by antitoxin binding at the cognate promoter.

Later work revealed that the ToxN toxin functions as a general RNase, and shares a conserved core fold with the type II toxins Kid, CcdB, and MazF[74,75]. Its RNase activity has two roles: firstly, it targets cellular mRNAs to disrupt translation, thereby inhibiting cell growth[74,76]. Its second purpose is to process the 5.5 repeat *toxI* transcript to generate 36 nucleotide ToxI-sRNA antitoxin oligomers[76]. These fold spontaneously as pseudoknots, with three ToxI-sRNAs and three ToxN protein monomers self-assembling to form an inactive heterohexameric ToxI-ToxN complex which occludes ToxN biochemical activity, thus neutralising toxicity[74,76,77] (Figure 1.3A).

Type III TA systems are an abundant and widespread class, located primarily on bacterial chromosomes, though with some type III loci identified on plasmids, prophages and potentially in archaea[75,78]. They can be divided into three families based on the toxin sequence similarity – *toxIN*, *cptIN* and *tenpIN* – which predominantly sit within the Firmicutes, Fusobacteria, and Proteobacteria phyla[78]. In addition to phage defence[31,73], these systems have also been implicated in plasmid maintenance[76,79,80], suggesting potentially diverse roles for type III systems in their host organisms.

1.1.4. Type IV TA systems

Similar to the type II, V, VI and VII TA systems, type IV antitoxins and toxins are proteins. However, the type IV TA systems are perhaps the most distinct class in that the antitoxin and toxin never directly (types II and VI), indirectly (type V), or transiently (type VII) interact. Instead, they target the same cellular process, but with antagonistic actions[81] (Figure 1.3B). In the first characterised type IV TA system, CbeA-CbtA, CbtA (cytoskeleton binding toxin) inhibits cell growth by physically interacting with the essential cytoskeletal proteins FtsZ and MreB. CbtA inhibits the GTPase activity and GTP-dependent polymerisation of FtsZ and ATP-dependent polymerisation of MreB *in vitro*, resulting in a lemon-shaped cell morphology[82]. The CbeA (cytoskeleton bundling-enhancing factor A) antitoxin acts antagonistically by directly interacting with FtsZ and MreB, enhancing the bundling of individual protofilaments and promoting the formation of higher-order FtsZ and MreB structures *in vitro*[81]. In this way, toxicity is neutralised by the antitoxin's opposing action on the same target as the toxin, with the antitoxin and toxin never interacting (Figure 1.3B).

The later characterised AbiEi-AbiEii type IV TA system from *Streptococcus agalactiae* (*S. agalactiae*) likewise has a non-interacting antitoxin and toxin protein pair, as demonstrated by co-immunoprecipitation studies in *E. coli*, which inhibit bacterial growth bacteriostatically

and reversibly[83]. Similar to the type III ToxI-ToxN system, AbiEi-AbiEii functions via an Abi mechanism to protect the bacterial population from bacteriophage predation, and can also promote stabilisation of MGEs[83]. The AbiEii toxin is a DUF1814 (domain of unknown function 1814) protein, which represents a family of uncharacterised putative nucleotidyltransferases (NTases) found widespread in approximately 3000 bacterial, archaeal and fungal genomes[83]. The COG5340 (cluster of orthologous group 5340) AbiEi antitoxin possesses an N-terminal winged helix-turn-helix (wHTH) DNA-binding domain and a C-terminal antitoxicity domain, both of which are important for transcriptional repression via negative autoregulation[84] (Figure 1.3B). This extra layer of transcriptional regulation is perhaps significant given the non-interacting nature of the type IV antitoxins and toxins.

1.1.5. Type V and VI TA systems

Two of the most recent TA classifications are the type V and type VI systems, with currently one characterised member apiece[85,86] (Figure 1.4). In the type V GhoS-GhoT system, the GhoT toxin is a small hydrophobic peptide that lyses membranes to arrest growth and, at high enough concentrations, creates nonviable "ghost" cells[86]. Ghost cells are formed following membrane disruption, with the cell centre becoming transparent and the cell poles darkening due to the presence of highly condensed intracellular material[2]. Expression of the GhoT toxin has previously been linked to increased persistence and early biofilm formation[86], suggesting an important biological role for cell survival.

The GhoS antitoxin differs to other protein antitoxin types in that it is a stable monomeric protein in solution, not subject to proteolytic degradation during stress, and which additionally does not bind its own promoter to repress transcription[86]. However, it primarily stands out due to its unique mechanism of action: GhoS is the first identified RNase antitoxin, which specifically degrades cognate *ghoT* mRNA *in vitro*, preventing translation of the toxin and thereby indirectly neutralising toxicity[86] (Figure 1.4A). Interestingly, GhoS shares structural homology and a conserved fold with several members of the CRISPR-associated (CAS2) family, which function as RNases to target specific single-stranded RNA[86,87]. This shared sequence-specific, post-transcriptional RNase activity suggests a potent tool which has evolved in different ways to benefit the cell, from controlling growth to countering phage invasion[86].

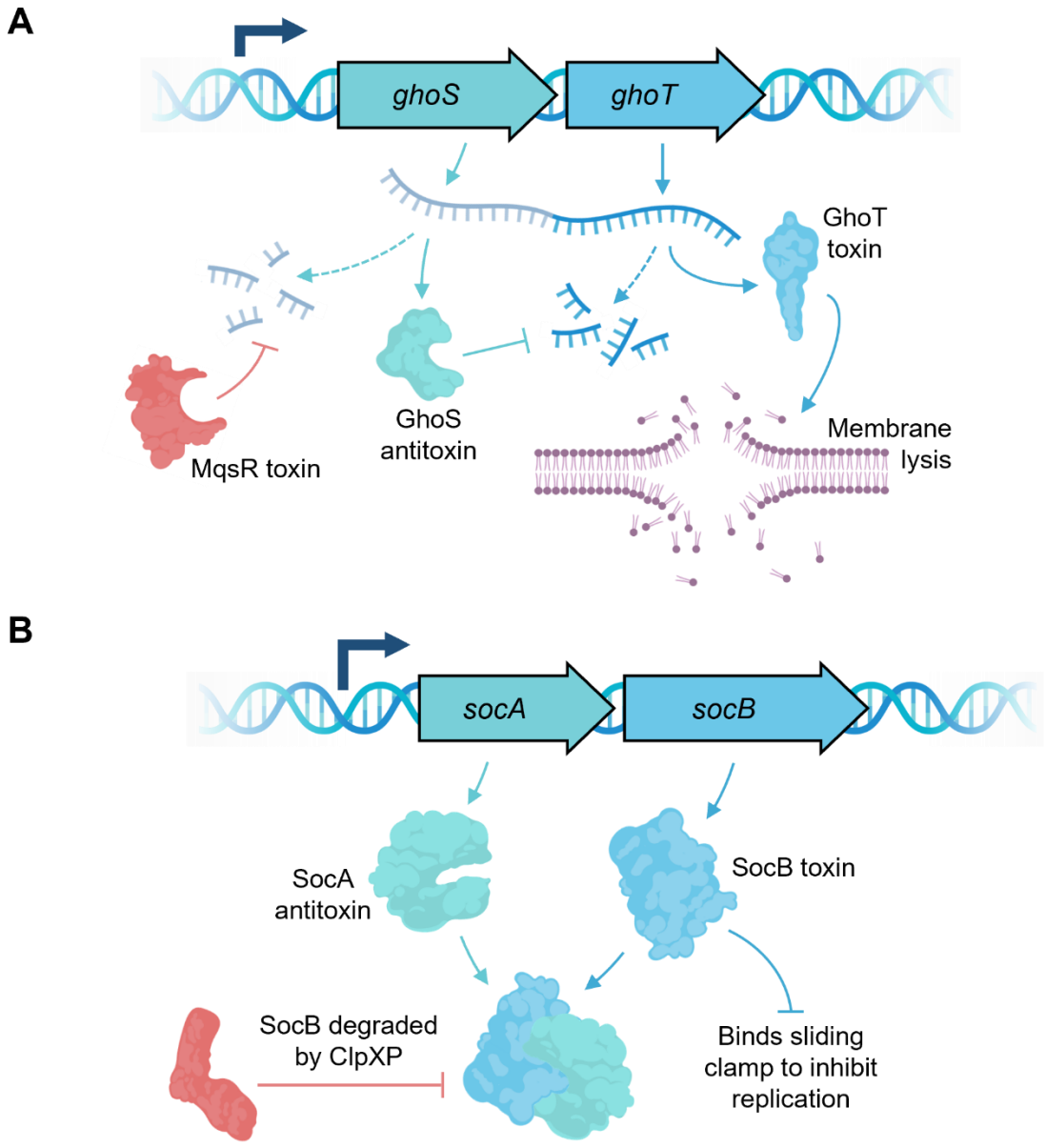


Figure 1.4. Schematic models of the GhoS-GhoT Type V and SocA-SocB Type VI TA systems. **(A)** The GhoS antitoxin is an RNase which specifically targets *ghoT* mRNA, thereby negating toxicity by preventing GhoT synthesis. Under cell stress conditions, *ghoS* mRNA is targeted and degraded by the MqsR toxin, facilitating GhoT translation and freeing the toxin to disrupt the cell membrane. **(B)** Both the SocA antitoxin and SocB toxin are proteins. SocA effects antitoxicity by binding the SocB toxin and mediating its interaction with, and degradation by, the cellular ClpXP protease. In the absence of antitoxin or protease, SocB binds the DnaN β-sliding clamp to inhibit DNA replication.

Another distinctive aspect of the GhoS-GhoT type V system is the involvement of a different TA system – type II MqsA-MqsR – in its regulation[88]. The MqsR toxin is a global RNase which degrades mRNAs featuring a 5'-GCU cleavage site[89]; the *ghoT* transcript lacks this cleavage site, however, *ghoS* mRNA contains three[88]. This results in the preferential degradation of *ghoS* mRNA and enrichment of *ghoT* mRNA in cells expressing MqsR, leading to GhoT toxin synthesis and growth arrest[88] (Figure 1.4A).

The type VI SocA-SocB TA system from *Caulobacter crescentus* is unique in that the SocA antitoxin alone is not sufficient for toxin neutralisation[85]. Rather, SocA acts as an adaptor protein with SocB as its substrate *in vitro*, tethering the relatively unstable toxin protein to the N domain of the cellular protease ClpXP, which in turn degrades SocB[85] (Figure 1.4B).

In the absence of either SocB or ClpXP, an accumulation of translated SocB in the cell results in growth arrest[85]. Bacterial DNA replication is carried out by DNA polymerase III (Pol III), which must associate with the β-sliding clamp, DnaN, to increase processivity[90]. In *E. coli*, DnaN interacts with a number of other proteins, including DNA Pols I, II, IV, and V[91], MutS and MutL, important for mismatch repair[92], and Hda, involved in replication regulation[93]. The SocB toxin binds DnaN via a conserved DnaN-binding motif, outcompeting the interactions of other β-sliding clamp-binding proteins[85]. This causes the collapse of the DNA replication fork and inhibition of replication elongation, which consequently prevents cell division and leads to growth arrest[85] (Figure 1.4B).

1.1.6. Type VII TA systems

The type VII classification was recently proposed as a means to better define the wide variety of biochemical functions among TA systems, and to prevent ambiguity within the type II classification[94]. The type II classification covers TA systems where toxin neutralisation is mediated via antitoxin binding. The type VII classification instead describes a similar yet distinct mechanism of toxin neutralisation, where rather than direct antitoxin-toxin binding, the antitoxin enzymatically inactivates the toxin via transient post-translational modifications (PTMs)[94] (Figure 1.5).

The type VII classification includes three currently characterised TA systems as examples: TomB-Hha[95,96], MntA-HepT[97], and $MenA_3$-$MenT_3$[98,99]. In each of these systems, the antitoxin neutralises toxicity by chemically modifying the cognate toxin, albeit in a distinct and specific manner. The *E. coli* TomB-Hha TA system has been reported as an important factor controlling biofilm formation[100], where the Hha toxin is a transcription regulator which

represses the transcription of regulatory rare tRNAs to inhibit type I fimbriae production[95]. Fimbriae production is a key step in biofilm formation; following the initial transient adherence of cells to a surface, fimbriae are a core component of the subsequent irreversible attachment of a subset of cells to form the initial biofilm monolayer[101]. Hha expression also exerts a toxic effect in the cell: the transcriptional repression of key tRNAs alters translation, causing the induction of lytic cryptic phage genes and general cellular proteases, such as ClpXP. This in turn is proposed to activate toxins of other TA systems via degradation of their cognate antitoxins[95]. A 2016 study found that the TomB-Hha TA system is oxygen-dependent, though the authors were unable to generate experimental data confirming a specific molecular mechanism of antitoxicity by the native *E. coli* TomB antitoxin[96]. However, they instead demonstrated that the *Yersinia enterocolitica* orthologue YmoB can functionally replace TomB to neutralise Hha toxicity[96]. They also showed that this neutralisation results from YmoB complexing with Hha, where it enhances the spontaneous oxidation of the conserved Hha residue C18 to a -$SO_{\chi}H$ (sulfenic, sulfinic or sulfonic acid)-containing species, thereby destabilising Hha structure[96].

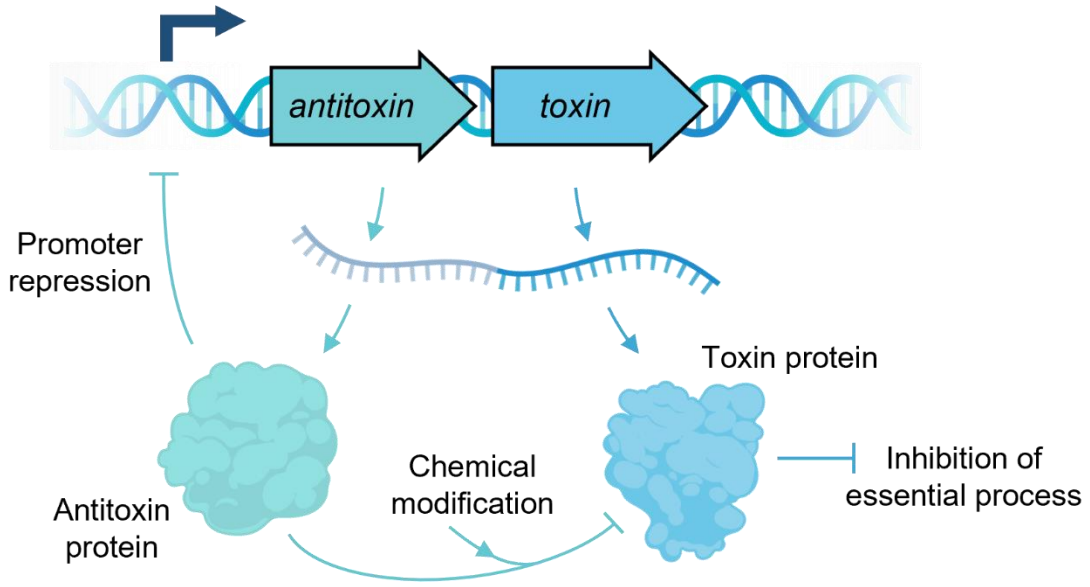


Figure 1.5. Schematic model of type VII TA systems. The antitoxin enzymatically neutralises cognate toxin activity via direct chemical modification of the toxin. This differs to the type II system, where inactivation occurs due to antitoxin-toxin binding; in the type VII system, the toxin instead acts as the antitoxin's substrate. In some instances, the antitoxin is additionally capable of regulating TA activity at the transcriptional level via promoter repression. Unmodified toxin is free to inhibit bacterial growth by targeting essential cellular processes.

The MNT-HEPN (minimal NTase domain; higher eukaryotes and prokaryotes nucleotide-binding domain) two-gene module is predicted to represent one of the most ubiquitous TA families in prokaryotes, with multiple putative biological roles[102]. *In vitro* studies of the MntA-HepT TA system from *Shewanella oneidensis* have shown that HepT (HEPN-domain toxin) is an RNase that cleaves mRNAs to inhibit cell growth[103]. HepT toxicity is neutralised enzymatically by MntA (MNT-domain antitoxin), which functions as an adenylyltransferase *in vitro*, polyadenylating residue Y104 of the HepT toxin, which sits adjacent to the HepT RNase motif[97]. The MntA antitoxin has also been shown to repress activity of the *mntA-hepT* promoter, proposed to bind two palindromes positioned near the -10 and -35 regions[104] (Figure 1.5). Notably, a bioinformatics analysis of the key MntA adenylyltransferase motif and HepT RNase motif identified their presence in a range of bacterial and archaeal species, potentially representing a widespread family of type VII TA systems[97].

In the case of the $MenA_3$-$MenT_3$ TA system (also reported under the name TakA-TglT), the $MenA_3$ antitoxin functions as a protein kinase that inactivates $MenT_3$ by phosphorylating an active site serine[98,99]. It should be noted that the functional, biochemical, and structural characterisation of the $MenA_3$-$MenT_3$ TA system is in part the subject of this dissertation and will therefore be discussed in greater detail later on.

1.2. Diverse toxin targets

As expected, given the ubiquity of TA systems among different prokaryotic organisms, the range of toxin targets is similarly diverse, with notable variety even within each TA type. Other than the variety, perhaps most significant is the essentiality of the targeted processes; the consequences of toxins targeting DNA replication or membrane integrity are potentially devastating to cell viability. However, disruption of certain toxin targets can be considered relatively less damaging to the cell than others[105]. For example, the effects of inhibiting translation, whilst more than sufficient to stall growth and reduce metabolic activity, can typically be reversed back to a normal physiological state, as opposed to the comparably more harmful effects of membrane lytic toxins or gyrase poisons[105]. This may account for the prevalence of toxins which inhibit protein synthesis, which might have been preferentially selected for due to their relatively less toxic nature[106]. Of additional interest, several toxin superfamilies have evolved which contain structurally related subfamilies with divergent mechanisms of action[9,65]. For example, the RelE ribosome-dependent mRNase family shares a structural fold with ParE gyrase inhibitors, whilst the MazF ribosome-independent mRNase

family is related to the CcdB gyrase poison[65]. Also notable is that certain functional classes of toxin appear more prevalent among certain TA types, such as membrane lytic peptide toxins, primarily encoded by type I families, and RNase toxins, which dominate the repertoire of type II systems[65]. A brief description of established toxin targets is given below, with examples of culpable toxins and their mechanism of action.

1.2.1. DNA replication

The unrelated CcdB and ParE type II toxins both inhibit DNA gyrase, a type II topoisomerase, to disrupt DNA replication and thereby inhibit growth. The originally characterised toxins are plasmid-encoded: CcdB is located on the F plasmid[28], and ParE is located on plasmid RK2[67,107,108]. CcdB takes a multifaceted approach to toxicity via its association with the dimerisation domain of the gyrase subunit GyrA[64]. In the absence of DNA, the interactions between CcdB and GyrA lead to inhibition of gyrase activity[26,109]. Alternatively, CcdB can also poison gyrase in a manner similar to quinolone antibiotics, by trapping it in a cleavage complex with cleaved DNA, preventing re-ligation and resulting in an accumulation of double-strand DNA breaks[27]. Finally, the CcdB-GyrA DNA-bound complex can also block the passage of RNA polymerase, thereby disrupting transcription[110]. ParE is functionally similar to CcdB, interfering with the activity of DNA gyrase to inhibit replication[67]. Accordingly, ParE2 from *Vibrio cholerae* (*V. cholerae*), a ParE homologue, also associates with GyrA to stabilise the gyrase-DNA cleavage complex[111]. However, the ParE/ParE2 mechanism has not been fully elucidated, and is thought to be different to that of CcdB as some CcdB-resistant mutants are still susceptible to ParE2 toxicity[111]. Another toxin, DarT from *M. tuberculosis*, can impair DNA replication via a completely different mechanism: DarT is an NTase which specifically and reversibly modifies single-stranded DNA (ssDNA) thymidines through adenosine 5'-diphosphate (ADP)–ribosylation[112,113]. Other mechanistically different toxins that target DNA include *E. coli* RalR, a type I DNase toxin which cleaves methylated and unmethylated DNA[58], and type II FicT, which adenylates DNA gyrase and topoisomerase IV in *E. coli* to block ATPase activity, thereby disrupting DNA topology[69].

1.2.2. Cell architecture

The *E. coli* type IV CbtA toxin disrupts cell division and elongation by directly targeting the cytoskeletal proteins FtsZ and MreB, inhibiting their polymerisation and causing a lemon-shaped morphology[81,82,114]. Two CbtA homologues are also present in *E. coli*: YkfI and YpjF,

from the respective YafW-YkfI and YfjZ-YpjF TA systems[114]. All three toxins are chromosomally encoded on separate cryptic prophages[114]. As with CbtA, YpjF interacts with both FtsZ and MreB, whilst YkfI interacts only with FtsZ[114]. These interactions occur at conserved residues on the FtsZ and MreB protein surfaces, disrupting polymerisation and the formation of important cytoskeletal filament architecture[114].

1.2.3. Membranes

Toxins targeting bacterial membranes are typically small, hydrophobic type I toxin peptides which usually contain an α-helical transmembrane domain, and sometimes a cytoplasmic or periplasmic domain[49,115]. Several type I toxins are understood to localise at cell membranes, where toxicity is effected by the destabilisation and deterioration of membrane integrity; this can cause membrane depolarisation and weakened proton-motive force (PMF), leading to a reduction in intracellular ATP concentration, and can also result in abnormal cell morphologies and ghost cell formation[2,30,53,55,116–120]. The downstream physiological consequences of toxin-mediated membrane damage can include a decrease in the rate of transcription, translation and replication, and degradation of cellular RNA[117,119,120]. In addition, although it does not target a bacterial membrane, the type II Zeta-like PezT toxin kinase from *Streptococcus pneumoniae* phosphorylates and inactivates important factors involved in peptidoglycan synthesis, which in turn inhibits cell wall synthesis and leads to cell lysis[121].

1.2.4. Protein synthesis

The most commonly reported mode of toxicity for type II toxins is inhibition of translation. Type II toxins commonly function as mRNases, disrupting translation and inhibiting growth by degrading protein transcripts. They typically fit into two categories: ribosome-independent mRNases and ribosome-dependent mRNases[122]. The first described ribosome-independent toxin was MazF from *E. coli*, which specifically cleaves the A▼CA sequence of free mRNAs[71]. Interestingly, a number of MazF homologues from other organisms have specificity for different length recognition sequences; either three, five, or seven bases[9]. As a result, the specificity and functional roles of MazF homologues are potentially highly diverse, dependent on the length of their mRNA recognition sequence, and their subsequent target promiscuity[71]. Contrastingly, the *E. coli* RelE mRNase toxin is ribosome-dependent,

associating with the ribosome A site to degrade bound mRNA, where it preferentially cleaves the UAG stop codon and obstructs translation elongation[123]. Another ribosome-dependent *E. coli* toxin, YoeB, instead interacts with the 50S ribosomal subunit of 70S ribosomes, where it blocks translation initiation by causing the cleavage of loaded mRNAs[124]. Furthermore, not all ribosome-dependent toxins function as RNases to disrupt translation. For example, the type II RatA toxin from *E. coli* blocks translation initiation by associating with free 50S ribosomal subunits, preventing their interaction with 30S subunits and consequently formation of 70S ribosomes[125].

Mechanisms for disrupting protein synthesis are not strictly limited to ribosome association and mRNA transcript degradation; different toxins target alternate elements of the translation pathway to block protein synthesis. In comparison to other members of the MazF toxin family, MazF-mt9 from *M. tuberculosis* is not a universal ACA-cleaving mRNase; instead, MazF-mt9 preferentially cleaves specific tRNAs in a sequence- and structure-dependent manner to impede protein synthesis and stall growth[126]. Separately, *M. tuberculosis* encodes up to 50 TA systems belonging to the type II VapB-VapC TA family[11,127]. Typically, VapC toxins are RNases, and the sheer number of familial VapC toxins encoded by *M. tuberculosis* suggests some functional divergence. Indeed, different *M. tuberculosis* VapC toxins have been shown to preferentially cleave different tRNAs[127,128], whilst *M. tuberculosis* VapC20 and VapC26 cleave the universally conserved Sarcin-Ricin loop (SRL) in 23S ribosomal RNA[127,129]. Some VapC targets are however conserved, as the VapCs from *Leptospira interrogans*, *Shigella flexneri* and *Salmonella enterica* serovar Typhimurium (*S*. Typhimurium) all cleave initiator $tRNA^{fMet}$ to impair translation[130,131]. Meanwhile, the GCN5-related *N*-acetyltransferase (GNAT)-fold toxins AtaT and TacT instead inactivate tRNAs via chemical modifications to disrupt translation, though with different target specificities. AtaT inhibits translation initiation by acetylating methionine-charged initiator $tRNA^{fMet}$ (Met-$tRNA^{fMet}$)[70], whereas the TacT toxins target translation elongation via acetylation of elongator aminoacyl-tRNAs[132,133]. The Doc toxin from bacteriophage P1 also blocks translation elongation by reversibly phosphorylating the EF-Tu elongation factor, thus preventing EF-Tu from catalysing the binding of aminoacyl tRNA to the ribosome A-site[68].

1.3. Physiological roles of bacterial TA systems

The locations of TA genes often provide clues as to the roles they play in the cell: plasmid-encoded TA modules such as CcdA-CcdB and ParD-ParE typically act as addiction modules, maintaining the presence of the plasmid in daughter cells following cell division via PSK[1,67], as described earlier (See, *1.1*; Figure 1.1). On the other hand, chromosomally encoded TA systems, which are found ubiquitously among prokaryotes, have been associated with numerous additional roles to addiction, with a common overarching theme being the bacterial response to stress[7]. Some bacteria carry a striking number of TA loci, with *E. coli* and *M. tuberculosis* encoding as many as 39 and 88, respectively[10–12]. Indeed, it is notable that bacteria encoding many TA modules tend to be slow growing or free-living, whereas those encoding no apparent TA loci, or very few, are obligate intracellular organisms[15]. Naturally, questions arise regarding why certain organisms possess such a relative wealth of TA modules, and for what reason? The benefit may be reflected in their diverse physiological roles, as well as in the persistent/chronic infections characteristic of pathogenic bacteria such as *M. tuberculosis*; different TA systems may provide condition-specific survival depending on the environmental context and a given stress[9,12]. A list of functional roles associated with TA systems is described below, alongside examples of responsible TA modules from a variety of organisms.

1.3.1. Abortive infection

Bacteriophages (phages) are the most abundant organisms on the planet, estimated to be upwards of 10^{31} in number[134]. The co-evolution of bacteria and phages has led to an evolutionary arms race, with bacteria developing an effective immune arsenal to resist the threat of infection, and phages likewise evolving an array of mechanisms to counter bacterial phage-resistance[135]. Bacterial defensive measures range from the well-established restriction modification systems, to the relatively more recently identified CRISPR-Cas systems[136]. Another bacterial phage-resistance mechanism is Abi, a bacterial innate immune response in which an infected cell induces growth arrest or commits altruistic suicide before phage replication can be completed[136]. This prevents phage maturation and propagation, and thereby protects the clonal bacterial population as a whole from further infection[65,136]. Abi is generally considered a last line of defence against phages. For example, where restriction modification targets invasive phage DNA early in infection, Abi is instead activated if phage replication reaches advanced stages, shutting down and/or killing the cell as a last resort[136].

Much like TA modules, Abi systems target a range of core cellular processes to impart toxicity, stall growth and induce cell death[136]. Indeed, there is considerable overlap with some TA toxin targets. For example, in response to T4 phage infection, the *E. coli* Lit protease shuts down translation to inhibit phage gene expression[137]. Separately, the *E. coli* Rex system imparts multi-phage resistance via the pore-forming RexB membrane peptide, which causes a loss of membrane integrity and potential, leading to a reduction of intracellular ATP[138]. Perhaps unsurprisingly given the mechanistic similarities of toxicity, a number of TA systems have been implicated in mediating an Abi phenotype (Figure 1.6).

The first TA system reported to be involved in phage inhibition was the type I Sok-Hok system on plasmid R1[32]. After infecting *E. coli* with T4 phage, the authors showed that overexpression of the Hok toxin resulted in strong T4 phage inhibition: cultured *E. coli* cells continued to grow, in comparison to the decreased optical density (representative of phage-induced cell lysis) of cells lacking Sok-Hok or with Hok expression repressed; efficiency of plating (EOP) and plaque size and number were reduced; phage replication and maturation time increased; and the latent period to cell lysis was delayed[32]. The reduced EOP and plaque count suggests that Hok overexpression prevented the production and propagation of T4 phage by causing cell death, consistent with an Abi phenotype[32]. Moreover, the continued growth of cultured T4-infected cells overexpressing Hok suggests that Hok expression protected the wider bacterial population from phage spread[32]. The authors proposed a model whereby infection with T4 phage and the resulting T4-induced transcription inhibition leads to the cessation of *sok* transcription. With no means to replenish it, the rapid degradation and loss of Sok antitoxin sRNA releases Hok, which in turn acts to arrest growth and kill the cell prior to the completion of phage replication[32] (Figure 1.6B).

More recent examples of functional Abi-TA modules include the type III ToxI-ToxN and type IV AbiEi-AbiEii TA systems, from *P. atrosepticum* and *S. agalactiae* respectively[31,83]. The ToxN toxin was identified based on sequence homology to the AbiQ Abi system from *Lactococcus lactis* (*L. lactis*), which itself was later discovered to belong to the type III AntiQ-AbiQ TA system[31,139,140]. ToxI-ToxN was shown to provide multi-phage resistance via a reversible, bacteriostatic TA/Abi mechanism, inhibiting growth through ToxN mRNase activity[31,73,74,76] (Figure 1.6B). Similarly to AbiQ, the AbiE system was originally recognised only as a lactococcal Abi system, where it acts late in the replication cycle, apparently restricting DNA packaging[141–143]. Among the identified Abi systems, AbiE was notably distinct in requiring the co-expression of two overlapping open reading frames (ORFs), designated *abiEi* and *abiEii*, to produce a phage-resistance phenotype[143]. Loci featuring two overlapping ORFs and one

upstream promoter are typical characteristics of TA systems[14–16]. Indeed, a later study found that the originally characterised AbiE module from *L. lactis* plasmid pNP40, as well as the chromosomal AbiE homologue from *S. agalactiae*, function through a type IV TA mechanism[83]. AbiEii toxicity is mechanistically distinct from ToxN/AbiQ, where it instead acts as a GTP-binding NTase to inhibit growth bacteriostatically, though with an as yet uncharacterised target[83]. Why these systems induce bacteriostasis to resist phage replication and killing is uncertain. Some have proposed that the effects of bacteriostasis and metabolic arrest eventually pass a point of no return that becomes bactericidal, leading to the altruistic cell suicide associated with Abi[144]. Others have suggested that Abi-/TA-induced bacteriostasis brings about a stationary phase-like growth state, which exerts a protective effect sufficient to disrupt phage replication, allowing the cell to recover when the phage threat dissipates[145].

1.3.2. Biofilm formation

Biofilms are complex, multi-cellular structures comprised of a physiologically heterogeneous mix of bacterial cells embedded within an extracellular matrix of self-produced polymers[146–148]. Biofilm cells are distinct from free-living, planktonic cells, presenting a range of genotypic and phenotypic differences which lend enhanced tolerance of and survival to stress and antibiotics, and contribute to recalcitrant infections[147–149].

The first reported TA system to be involved in biofilm formation was the *E. coli* type II MqsA-MqsR system, where gene expression profiling revealed that the *mqsR* toxin gene is induced in biofilm cells[89,150]. In a 2006 study, Barrios *et al*. reported that deletion of *mqsR* leads to diminished motility and biofilm formation relative to a Wild-Type (WT) strain, which can be restored upon MqsR expression *in trans*[151]. The *mqsR* deletion also results in the differential expression of numerous genes involved in cell motility and biofilm formation, including down-regulation of flagella and curli regulator genes, implicating the MqsR toxin as a global regulator of biofilm-forming pathways[151]. Barrios *et al*. also reported that the biofilm-forming activity of the quorum-sensing signal autoinducer 2 (AI-2) is mediated by MqsR[151]. They found that AI-2 stimulates expression of MqsR, which consequently triggers a regulatory cascade that promotes biofilm formation[151].

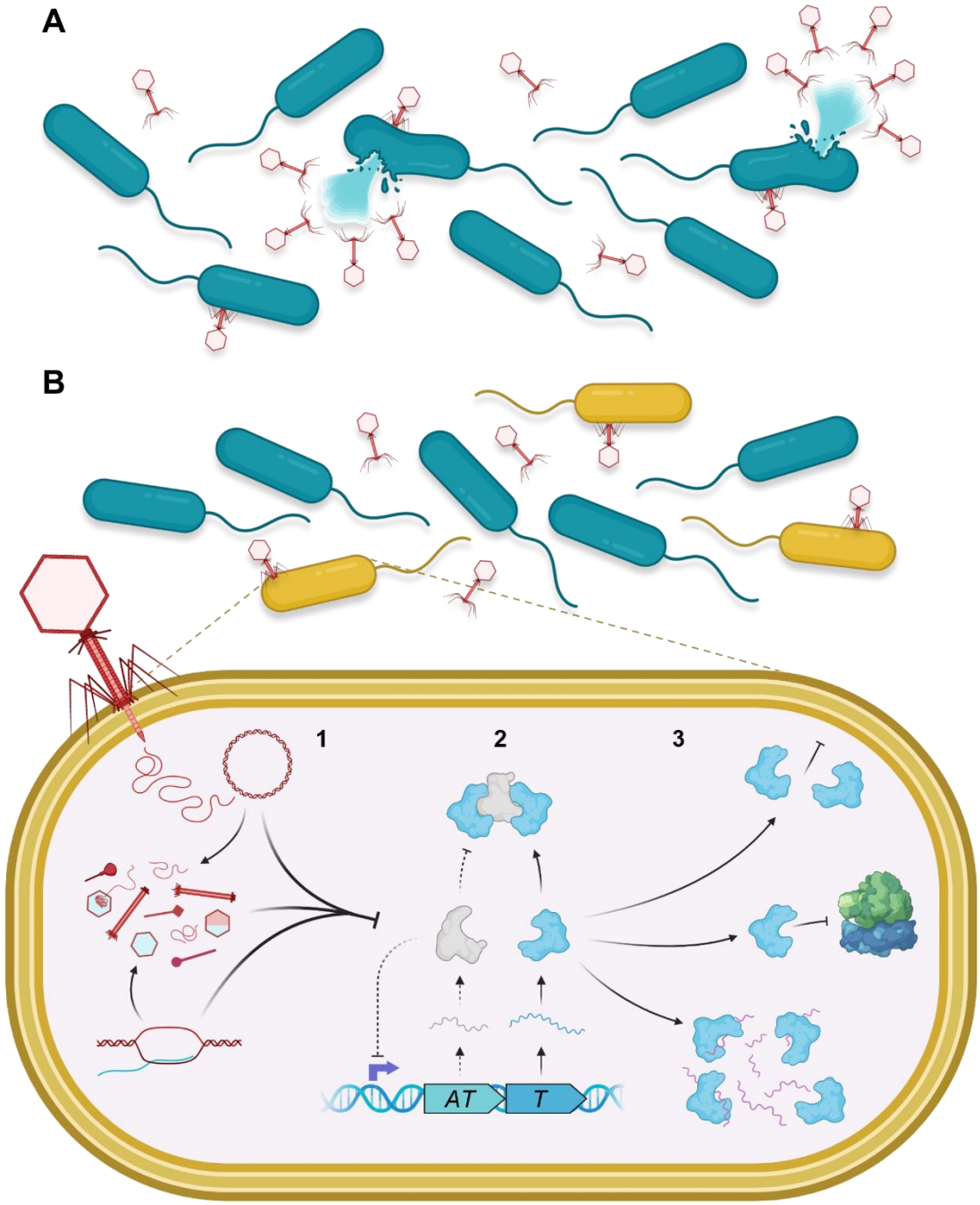


Figure 1.6. Model of TA-mediated Abi phage defence. (A) Bacterial cells (teal) of a clonal population lacking Abi protection are vulnerable to phage infection, replication, and propagation. **(B)** In the TA/Abi model, phage-infected cells (yellow) protect the population from viral spread through TA-mediated Abi phage defence. (1) Phage infection hijacks cellular machinery, repurposing it for phage DNA replication and expression of phage proteins. (2) Phage-mediated translational shutdown triggers a regulatory cascade which causes antitoxin (AT) degradation by cellular proteases. By hijacking the host gene expression apparatus, phage infection impairs antitoxin synthesis, which cannot be replenished when degraded. Consequently, antitoxins are unable to inhibit toxin (T) activity as normal. (3) Toxins, freed from the neutralising action of cognate antitoxins, target essential cellular processes, for example cleaving cellular mRNAs, inhibiting translation, or targeting the bacterial membrane. This activity stalls growth, preventing phage replication and spread to neighbouring cells.

Moreover, the cognate MqsA antitoxin, which functions as a global regulator alongside its antitoxic properties[152], also features in biofilm regulation. MqsA binds two palindromes of the *mqsA-mqsR* promoter to negatively regulate its own expression[89]; however, MqsA can also bind the promoters of other genes important for bacterial physiology and stress response[153,154]. Significantly, MqsA can bind *mqsA-mqsR*-like palindromes in the promoter of *csgD* (P_{csgD}), which encodes a master regulator controlling curli production and biofilm formation[155]. MqsA also binds and represses the promoter of the master response regulator *rpoS* (P_{rpoS}), which, among other roles, induces *csgD* expression as well as regulating genes responsible for the synthesis of the biofilm-stimulating 3,5-cyclic diguanylic acid (c-di-GMP) second messenger[152]. Therefore, MqsA-P_{csgD} and MqsA-P_{rpoS} binding represses transcription of the respective regulons, inhibiting curli and c-di-GMP production and impairing biofilm formation[152,155].

Other examples of TA systems influencing biofilm formation include the type VII TomB-Hha TA system (See, *1.1.6*), which has been found upregulated in *E. coli* biofilms compared to planktonic cells[150]. Deletion of this complete TA module resulted in a reduction in biofilm mass, as well as decreased conjugation and cell aggregation, indications of deficient biofilm formation[100]. Closer scrutiny of the mechanistic detail revealed that the Hha toxin inhibited biofilm formation by restricting the transcription and availability of rare codon tRNAs, which in turn repressed production of type I fimbrae[95]. Hha expression also caused biofilm dispersal and cell lysis, the latter via activation of cellular proteases – proposed to consequently degrade the antitoxins of other TA systems, thereby freeing their toxins – and by activating cryptic prophage lytic genes[95]. Expression of TomB reduced Hha toxicity when tested in a two-plasmid system, and presumably serves as a counterbalance to Hha to regulate WT biofilm formation *in vivo*[95]. Interestingly, the influence of other TA systems in Hha-biofilm regulation was corroborated in a separate study by Kim *et al*., who reported that a strain of *E. coli* MG1655 with five TA systems deleted abolished Hha-mediated biofilm dispersal[156]. Furthermore, the type V GhoS-GhoT TA module has also been implicated in biofilm regulation, though in contrast to the biofilm impairing effects of Hha, the GhoT toxin seemingly promotes biofilm formation; deletion of *ghoT* decreases biofilm formation, whereas loss of the *ghoS* antitoxin gene increases biofilm formation[86].

1.3.3. Stress tolerance and cell survival

Although increasingly challenged in recent years, TA systems have long been implicated in the bacterial response to stress. Various conditions pose a threat to bacterial populations and cell viability, for example nutrient starvation, DNA damage, antibiotic exposure, and heat shock, to name but a few[157,158]. Many examples of stress-responsive TA systems implicate their activity with the transition to a dormant, metabolically slow growth state, or persister state[7,34,159–161]. This physiological state lends high stress tolerance and enhanced survival to the cell in question; by inhibiting core cell functions and slowing growth, the cell is proposed to wait out the otherwise harmful pressure until it abates[34,162]. However, persistence is only briefly referred to in this section where relevant. TA systems and persistence are instead discussed in more detail in section *1.4* below, as the contentious nature of TA involvement merits greater attention. Examples of various cell stresses are given below, alongside associated TA systems that are implicated in mediating cell survival.

DNA damage can prove deleterious to bacteria if not met with effective countermeasures. It can have numerous causes: environmental factors, such as UV light and ionising radiation, as well as chemical inducers, for example cross-linking or oxidising agents[163]. Other sources of DNA damage can include antibiotics, such as fluoroquinolones, which poison DNA cleavage complexes, whilst intracellular metabolism can generate reactive oxygen species (ROS)[163,164]. DNA damage is of significant concern to the cell, as detrimental mutations may impair fitness and lead to cell death. The ability to react to and repair DNA damage is therefore essential for cell viability: as a result, bacteria have evolved the SOS response, a global regulatory network that is induced in response to DNA damage[163]. The SOS response is tightly regulated by the key SOS effector proteins, LexA and RecA[165]. During normal growth, dimeric LexA represses transcription of the SOS regulon by binding a specific site (the SOS box) in the respective gene promoters, interfering with RNA polymerase activity[166]. However, in response to DNA damage, RecA binds ssDNA and an NTP as cofactors to enter an active state (RecA*)[163]. RecA* triggers self-cleavage of LexA and exposes key residues which are targeted by the housekeeping proteases ClpXP and Lon, leading to degradation and reduction of the cellular pool of LexA, and derepression of SOS genes[163]. The SOS regulon comprises a range of genes which fulfil different functions to counteract DNA damage, including homologous recombination, excision repair, and inhibiting cell division[167,168].

Notably, TA systems have also been reported to be activated under SOS-inducing conditions, suggesting that certain TA modules may expand the SOS regulon. The type I *istR-tisAB* TA

locus encodes the toxic membrane peptide TisB and the antitoxin antisense sRNA IstR-1, and was the first TA system to be identified as activated by the SOS response[54,169]. The promoter of *istR-tisAB* contains an SOS box, suggesting that SOS-induced cleavage of LexA and derepression of *istR-tisAB* might induce transcription and synthesis of TisB, resulting in slowed growth[54,167]. Indeed, induction of the SOS response depletes the cellular stock of IstR-1 antitoxin, leading to an accumulation of *tisAB* mRNA, enhanced toxicity, and growth arrest[54]. This is suggested to form a line of defence against DNA-damaging agents such as antibiotics: the activity of the membrane peptide TisB in lowering PMF and decreasing cellular ATP induces an inert growth state that reduces the availability and/or activity of antibiotic targets and encourages survival[169]. Affirming this, the fluoroquinolone ciprofloxacin has been shown to activate the *tisAB* promoter and induce multi-drug tolerant (MDT) persister formation in a *tisAB*-dependent manner[169]. A number of other SOS-activated TA systems have also been reported, including type II DinJ-YafQ (albeit with contradictory reports of LexA regulation[167,170]) and YafN-YafO, and the type I SymR-SymE, AgrB-DinQ, and SokE-HokE systems[118,167,168]. Interestingly, in the case of the YafN-YafO and SymR-SymE systems, despite upregulation following SOS induction, no SOS-induced TA-dependent phenotypes, such as enhanced cell survival, have so far been reported[57,171].

Another significant cell stress is nutrient starvation. The intracellular signalling alarmone (p)ppGpp (used generally to refer to guanosine tetraphosphate and guanosine pentaphosphate[172]) functions as a global regulator controlling the expression of genes involved in core cellular processes and virulence[173–176]. Synthesis of (p)ppGpp typically occurs via two pathways in *E. coli*: SpoT-dependent and RelA-dependent, although in Gram-positive organisms only one bi-functional SpoT/RelA homologue (Rel) is responsible[173,177]. The latter pathway, involving the stress response regulator RelA, is responsible for mediating the stringent response (SR), the bacterial response to amino acid (AA) starvation[173]. RelA associates with the ribosome, where its (p)ppGpp-synthetase activity is activated in response to uncharged tRNAs in the ribosomal A site; a symptom of limited AA availability or impaired tRNA aminoacylation[173]. Accumulation of (p)ppGpp via the SR is associated with reduced mRNA synthesis and transcriptomic changes affecting core metabolic processes; these changes are proposed to redirect transcriptional resources to genes essential for starvation survival, whilst de-prioritising genes required for growth[173,176,178,179]. Interestingly, the SR also induces expression of TA modules, consistent with a need to lower metabolic burden under stressful AA starvation conditions[178]. Indeed, numerous reports have associated the activation of TA systems with the SR. Accumulation of (p)ppGpp activates Obg, a conserved

GTPase associated with the control of key cellular processes[180]. Obg induces expression of the type I HokB toxin, which mediates survival via membrane depolarisation, growth arrest and induction of the persister state[180].

Other examples of SR-activated TA systems include *E. coli* RelB-RelE and the homologues YafN-YafO, HigA-HigB, and YgiT-YgiU, which encode mRNase toxins[181,182]. Interestingly, *yafN-yafO* is also upregulated by the SOS response. However, the *yafN-yafO* locus is controlled via two promoters: one, located within its own operon and autoregulated by YafN, and the second LexA-regulated promoter belonging to the *dinB* operon immediately upstream[181]. The mRNase activity of activated YafO may therefore be beneficial in response to DNA damage, by reducing the metabolic load and slowing growth, to allow DNA repair mechanisms to act[171,181]. A further example of an SR-activated TA module is the type II *hicB-hicA* system, transcription of which was strongly induced after the addition of serine hydroxymate (SHX) to exponentially growing cells, which replicates nutrient starvation and activates the SR[183,184]. Interestingly, this system also responded to exposure to two other stresses: the addition of α-methyl glucoside to cells growing in minimal medium, inducing carbon starvation, and addition of chloramphenicol to disrupt translation, both stimulated *hicB-hicA* transcription[183]. HicA is an RNase toxin, and as with RelE, YafO, HigB and YgiU mentioned above, their respective expression reduces the rate of global translation[181–183]. Toxin expression in response to AA starvation may therefore benefit cells; by reducing the amount and availability of uncharged tRNAs to enhance translational fidelity, as well as promoting better responsiveness and adaptability to fluctuating nutrient levels[183]. A number of other TA systems have also been reported to be stimulated by multiple cell stresses. By way of TA promoter-*gfp* fusion assays, Shan *et al*. showed that a variety of stresses including isoleucine starvation, osmotic stress, phosphate stress, acid stress, and NaCl stress, at physiologically relevant levels, and to varying degrees of efficacy, could upregulate expression of ten type II TA systems[185]. Transcription of type II TA systems is autoregulated by binding of the antitoxin or antitoxin-toxin complexes to the cognate promoter. Therefore, Shan *et al*. reasoned that the observed promoter activity reflected a reduction in antitoxin levels compared to toxin levels, and therefore stress-activated toxin[185].

An additional role for TA systems in the SR has also been proposed. Within a few minutes of AA starvation, (p)ppGpp levels peak, which is soon followed by a decrease proposed to be linked to (p)ppGpp-mediated reduction of cellular mRNA[179,186–191]. This forms a negative feedback loop, where the decreased transcription of mRNA limits the number of empty ribosomal A sites, reducing RelA activity and returning the cell to a pre-starvation state[187].

Tian *et al*. suggested that specific TA modules function as part of this feedback mechanism to further regulate levels of (p)ppGpp and curtail the SR[187]. They revealed that these TA systems are rapidly activated early in the SR after AA starvation, possibly a result of (p)ppGpp-mediated protease activation and consequent antitoxin degradation[187]. The activated mRNase toxins help reduce the mRNA pool, thereby negatively regulating (p)ppGpp and potentially promoting a return to pre-starvation levels and a regular growth state[187].

A final example of a stress-linked TA system is MqsA-MqsR, which has previously been strongly implicated in the bacterial response to a variety of cell stresses. For instance, the toxin MqsR, an mRNase, enhances resistance to the bile acid deoxycholate by degrading *ygiS* mRNA[192]. YgiS is a periplasmic protein which increases intracellular uptake of deoxycholate, causing membrane leakage and impaired growth[192]. It was suggested that under conditions of bile acid stress, degradation of the neutralising MqsA antitoxin by Lon releases MqsR to inhibit YgiS synthesis and promote bile acid resistance[152,192]. Transcriptional regulation of *mqsA-mqsR* has also been linked to other stresses, including nitrogen starvation and heat shock in *E. coli*, and copper stress in the phytopathogen *Xylella fastidiosa*[193–196]. In addition, the antitoxin MqsA, which can regulate multiple promoters besides its own, also represses expression of the RNA polymerase sigma (σ) factor RpoS[152,197]. In *E. coli*, RpoS functions as a general stress regulator, in contrast to more specific and adaptive responses such as the SOS or SR[197]. Various stresses can trigger an accumulation of RpoS, which subsequently induces expression of a broad and generalised stress resistance regulon[197]. During stress, degradation of MqsA by Lon protease derepresses *rpoS*, leading to the RpoS-regulated general stress response[152]. Repression by MqsA under stress-free growth conditions would therefore appear important, by preventing premature and detrimental *rpoS* induction.

Nevertheless, the link between TA systems and stress response has become increasingly challenged, with several studies presenting counterevidence to dispute the role. Primarily, the point of contention surrounds the limited evidence for physiological changes mediated by TA systems under stress, contradictory data, and the inability to reproduce certain results[21,198]. For example, Fraikin *et al*. recently reported that exposure of *E. coli* to a range of stress conditions, such as bile salts, oxidative stress, or AA starvation, does not regulate transcription of *mqsA-mqsR*[199]. They found that MqsA did not regulate the *rpoS* promoter, and noted that both overexpression of *mqsA* and deletion of *mqsA-mqsR* produces no stress resistance phenotype, with oxidative stress and bile acid sensitivity, as well as ability to form biofilms, comparable to a WT strain[199]. Similarly, a prior study found that an *E. coli* strain deleted of five mRNase-encoding TA systems exhibited no survival or fitness defects

compared to a WT strain after exposure to stress[200]. In addition, a recent report by LeRoux *et al*. showed that while diverse stress conditions can activate transcription of TA modules, stress was not a determining factor in toxin activation[21]. They found that an *E. coli* MG1655 strain deleted of ten chromosomal type II TA systems suffered no growth defects during and after exposure to various stresses, and that stress did not result in enhanced RNA cleavage by the RNase toxins[21].

The conflicting evidence of recent studies serves to highlight our still limited understanding of how, or if, stress contributes to TA system regulation and activation[21,198]. Indeed, the increasing debate certainly raises questions as to their functional role and/or essentiality in stress response[21,198]. Why, for example, would vital components of bacterial stress response and survival be arbitrarily distributed among bacterial genomes[198]? Such questions might imply that the primary function of TA systems is rather linked to their genetic location, encoded as they often are on MGEs[198]. Whilst there is a wealth of existing literature evidencing a link between TA modules and stress, this does not necessitate dogmatic adherence to the proposed models of past studies. Equally, however, the breadth and scale of this literature is compelling, and lends credible weight to the idea that specific systems indeed function as components of bacterial stress response pathways.

1.3.4. Addiction and anti-addiction: stabilisation of chromosomal elements

TA systems were originally found on plasmids, involved in their maintenance via PSK[1,67] (Figure 1.1). The interdependence of toxins and antitoxins is vital to this role; normal growth and cell survival requires the presence of both the antitoxin and toxin to negate toxicity. As a result, TA modules have gained the moniker of “addiction molecules”, because bacterial offspring are necessarily addicted to the presence of the antitoxin, and therefore the TA locus[201].

Chromosomal TA modules are mostly located in the accessory genome, often alongside virulence and antimicrobial resistance (AMR) genes, found on MGEs such as cryptic prophages, super-integrons (SIs), transposons, and integrative and conjugative elements (ICEs), the product of horizontal transfer[106,202]. Type I systems appear to be an exception, with evidence suggesting that these systems have developed via lineage-specific duplication[45]. For the majority of chromosomal TA systems, however, their heterogeneous distribution is strongly suggestive of a role for horizontal gene transfer in their dissemination amongst bacterial species. For example, the *E. coli* O157:H7 chromosomal *ccdA-ccdB* TA system is

encoded between the *folA* and *apaH* genes[203]. However, this system is not consistently distributed among all *E. coli* species[203]. In those strains lacking *ccdA-ccdB*, the *folA-apaH* intergenic region contains a palindromic sequence which happens to be a target of insertion by the IS1397 MGE[203,204]. This might indicate a historical transposition of the *ccdA-ccdB* locus into this transposable site, and provides a compelling example of genome hopping by TA modules[203].

The selfish nature of TA systems hints at how these loci maintain themselves within the accessory genome, via mechanisms reminiscent of PSK, as well as the role they play in maintaining their typically unstable host MGEs. SIs, for example, are large chromosomal elements containing variable arrays of gene cassettes, each of which is flanked by an *attC* recombination site that is targeted by the integron-encoded integrase[205–209]. Vibrionaceae SIs can be in excess of 120 kb in length, featuring a large number of genes and associated recombination sites[210,211]; as such, instability of this element would be expected[6]. Despite this, SIs are typically relatively stable, determined by selective pressure and/or pre-existing stability-determining mechanisms[6]. Corresponding with the latter point, SIs commonly encode TA loci among their many gene cassettes, where their presence stabilises SIs and prevents large-scale deletion events[6,212]. The *relB-relE* and *parD-parE* systems encoded on the *Vibrio vulnificus* SI suppress mass deletion events otherwise mediated by the integron integrase[6]. In addition, both *higA-higB* TA loci from the *V. cholerae* SI on chromosome II can stabilise an unstable test plasmid in *E. coli*, raising the prospect of a role in SI maintenance[213]. TA modules on SIs are thus proposed to contribute to streamlining of the genome via a process of “microevolution” over “macroevolution”; by maintaining the general genomic region and preventing mass deletions through a PSK mechanism, but still permitting integrase-mediated site-specific rearrangements of individual gene cassettes[6,212]. Other examples of MGE-stabilising TA modules include the widely distributed *sgiA-sgiT* TA system, first found encoded on the multi-drug resistance (MDR) *Salmonella* genomic island 1 (SGI1), which stabilises and preferentially maintains SGI1 in the presence of an incompatible plasmid[214]. SXT, a MDR ICE found in *V. cholerae* clinical isolates, is maintained in daughter cells following excision to its transient extrachromosomal form by excision-mediated upregulation of the SXT-encoded *mosA-mosT* TA module[215]. The *E. coli relB-relE* TA locus is encoded in the Qin cryptic prophage, where it seemingly also confers a stabilising effect[200]; when cloned into an unstable test plasmid, *relB-relE* provides stabilisation[3]. Interestingly, the presence of TA loci on plasmids and other MGEs carrying MDR genes, and their stabilising

effect on these elements, indicates a role for TA systems in maintaining and disseminating MDR genes amongst bacteria[216].

Interplay between chromosomal- and plasmid-encoded TA systems has also been observed. In the case of the homologous *chpBI-chpBK* and *kis-kid* systems, the chromosomal ChpBI antitoxin has been shown to neutralise toxicity of the Kid plasmid toxin[217]. Therefore, in addition to stabilising MGEs, the cross-interactivity of chromosomal TA systems with plasmid-borne homologues may also be a major contributor to bacterial evolution; either by facilitating the acquisition of toxin-only plasmids, and the genetic material therein, or increasing fitness by promoting the loss of plasmids with homologous TA modules[212]. Regarding the latter point, some chromosomal TA modules have been proposed to function as anti-addiction systems. In this model, chromosomally-encoded antitoxins act as anti-addiction molecules which can neutralise the toxicity of plasmid-borne toxins with sufficient homology[201]. This then negates plasmid-mediated PSK and allows plasmid loss without cell death, which can help to increase cell fitness and fix the chromosomal TA system in the population[201]. In turn, this is proposed to drive the evolution and selection of plasmid-borne TA modules which can no longer be neutralised by the anti-addiction elements[201]. Supporting this model, De Bast and colleagues showed that the antitoxin from the *Erwinia chrysanthemi* (*E. chrysanthemi*) 3937 *ccdA-ccdB* system (ccd_{Ech}), when tested in *E. coli*, is able to inhibit the toxin homologue from the *E. coli* F plasmid (ccd_F)[218]. Furthermore, when integrated into the *E. coli* MG1655 chromosome, ccd_{Ech} effectively negated PSK and maintained cell viability despite the presence of plasmid-borne ccd_F, as well as providing a fitness advantage to the host cell under ccd_F-mediated PSK conditions[218].

An additional alternative hypothesis is that TA systems are simply selfish elements; ultimate evolutionary survivors which migrate between genomes on plasmids and chromosomal elements by horizontal gene transfer, all the while enforcing their maintenance due to their inherently selfish nature[106,201]. As such, the observed stabilisation of genetic elements by chromosomal TA modules may simply be a result of their addictive properties and the random location they migrate to[106,201].

The uniformity of this stabilising role among chromosomal TA modules is however questionable. Different systems in different organisms appear to provide varying levels of stability, while the stabilising effect of other chromosomal TA systems has been shown to be null. Pathogenic *E. coli* O157:H7 harbours both a chromosomal *ccdA-ccdB* TA system (ccd_{O157}), as well as a *ccdA-ccdB* module encoded on an F-related pO157 virulence plasmid (ccd_{pO157-}

$_F$)[203]. Interestingly, when cloned into an unstable plasmid, *ccd*$_{O157}$ is unable to mediate PSK, whereas *ccd*$_{pO157\text{-}F}$ can mediate PSK even in the presence of its chromosomal counterpart[203]. Furthermore, unlike CcdA$_{Ech}$ (described earlier), CcdA$_{O157}$ is unable to counteract CcdB$_{pO157\text{-}F}$ toxicity; however, CcdA$_{pO157\text{-}F}$ can neutralise the toxicity of chromosomal CcdB$_{O157}$[203]. This would heavily favour PSK by plasmid-encoded *ccdA-ccdB*, ensuring its stability and coexistence, and suggesting that some chromosomal homologues of plasmid-borne TA systems are adapted to serve an alternative function. In line with this, it has also been suggested that, following integration into the host genome, the ability of these TA systems to confer stability may fade over time; that they evolve new biological functions to suit their new genomic location[200,203].

1.4. Bacterial persistence and TA systems: a rocky relationship

Many bacteria can cause long-term infections in their hosts, with *M. tuberculosis* one of the most well-known pathogens associated with latent, chronic infections[219,220]. Bacteria use a range of survival strategies to establish and maintain infections, such as by evading the host immune response or withstanding antibiotic treatment[157,158,161]. The latter strategy can be accomplished in several distinct ways; via drug "resistance", "tolerance", or "persistence"[221,222].

While all three strategies typically promote bacterial survival under antibiotic pressure, they are markedly different. Resistance arises from heritable gene mutations, occurring either randomly or due to drug exposure, which affect different mechanisms and processes, for example efflux pumps or modifications to a drug target, rendering the drug ineffective[223]. Drug-resistant bacteria continue to grow during treatment while susceptible bacterial are killed off; the level of resistance is quantified by the minimum inhibitory concentration (MIC) of a particular drug required to prevent bacterial growth[223]. In other words, a higher concentration is needed to confer the same antibiotic effect on a resistant strain as a susceptible strain[221].

In contrast, tolerance describes whole bacterial populations that are able to withstand transient exposure to high, otherwise lethal drug concentrations, without a change in MIC, by transitioning to a dormant growth state[221,222]. In some cases, dormancy may be an inherited feature; alternatively, it can be induced by growth-limiting environmental conditions, or by bacterial responses to specific antibiotics[221]. This slow growth state can impair cellular processes such as translation or cell wall synthesis; consequently, this reduces

the ability of bactericidal antibiotics such as β-lactams to kill the cell, as these drugs target active cell wall assembly[221]. The drug concentration becomes essentially redundant, due to the scarcity of its target, meaning a longer treatment duration is instead required to kill the tolerant bacteria[221].

Persistence is a similar yet distinct mechanism. Persister cells are generally described as small, clonal subpopulations of bacterial cells that have undergone a spontaneous, transient, and reversible switch to a persister phenotype, characterised by slowed or stalled growth, and low metabolic burden[33,224,225]. This persister state renders the cell highly stress-tolerant and able to "persist" in the face of external pressures, prolonging the treatment duration these cells can survive, while the wider bacterial population is rapidly killed[33,161,226]. When these stresses are lost, the persister subpopulations "resuscitate" and revert to a normal growth state to repopulate their niche, leading to recalcitrant infection[33,161,226]. Persistence is distinct from antibiotic resistance, with the former a transient, reversible and nonheritable phenotypic shift, and the latter arising from gene mutations which generate heritable resistance and render antibiotics ineffective[33,227]. However, the two are fundamentally linked; the increased exposure of cells to antibiotics during persistent infections may lead to a greater likelihood of genetic resistance developing as a result[161,228,229].

It has been suggested that bacterial populations use stochastically occurring persister cells as an "insurance policy"; whilst the presence of small metabolically-inactive persister populations may lower overall population fitness, these persister cells can save the population from complete eradication when exposed to sudden stress[226,230]. In contrast, others have argued that the idea of spontaneously generated persister subpopulations is a misconception[10], as various environmental factors have been shown to influence conversion of normal-growing cells to the persister state[10,169,231–235], and that a more likely physiological origin for persistence is stress response[10]. TA systems were therefore considered ideal proponents of persistence, given their established role in modulating bacterial growth dynamics in response to stress[33,34,236]. In the simplified TA persistence model[7] (Figure 1.7), exposure of a bacterial population to external stress leads to antitoxin proteolysis and toxin activation in a subset of cells. Toxin-mediated growth inhibition induces the switch to a stress-tolerant persister state, which enables cell survival under otherwise lethal conditions. This presents through biphasic killing kinetics, where following the initial loss of susceptible cells, the gradual plateauing of the killing curve reveals the presence of stress-tolerant persisters[34,221] (Figure 1.7B). Following the relief of stress, antitoxins are replenished and subsequently neutralise toxicity; this allows cells to revert to a normal growth state in order

to repopulate the niche[7]. Numerous studies have provided evidence linking specific TA systems to persister formation, with this role becoming arguably the most contentious associated with these systems to date[10,237,238].

The first identified genetic component affecting persistence was *hipA* (high persister protein A)[239], of the type II HipB-HipA TA system[240]. The HipA toxin was first implicated in persistence nearly four decades ago when a mutant variant, HipA7, was isolated from *E. coli* pre-treated with ampicillin[239]. The HipA7 mutant leads to an increase in persister frequency, resulting in increased survival to ampicillin treatment[224,239]. The HipA toxin functions as a serine-threonine kinase which phosphorylates GltX, an aminoacyl $tRNA^{glu}$ synthetase, at Ser239 of its ATP-binding site to inactivate it[241,242]. Interestingly, HipA possesses additional substrates to HipA7, which only targets GltX; this is suggested to account for the increased toxicity of HipA compared to HipA7[240,243,244]. Despite differences in both substrates and toxicity, both HipA and HipA7 have been shown to induce persistence, suggesting HipA-/HipA7-mediated GltX phosphorylation is a sufficient mechanism for persister formation in *E. coli*[241,242,244]. In this model of persister formation, the phosphorylation of GltX by HipA causes an accumulation of "hungry" uncharged $tRNA^{glu}$ in the cell[241,242]. The presence of hungry $tRNA^{glu}$ triggers activation of the SR regulator RelA and production of (p)ppGpp, which in turn activates the SR to cause persistence[241,242].

Since the identification of HipA as a persister-associated toxin, the number of reports linking TA systems to persistence has grown substantially. Expression profiles from *E. coli* persister cells, obtained from *E. coli hipA7* and WT strains, reveal the differential expression of several TA genes[225,245]. Multiple research groups have also demonstrated that expression of type II toxins leads to persister phenotypes in *E. coli*[245–248]. Commonly, type II toxins are mRNases, suggesting a link between inhibition of protein synthesis, which leads to stalled growth, and the onset of persistence. For example, the controlled expression of the MazF and RelE type II mRNase toxins causes MDT phenotypes and increased persistence[245,246]. Similarly, overexpression of the YafQ toxin increases MDT in biofilm cells, whilst deletion of chromosomal *yafQ* decreases MDT[247]. MazF also induces an MDT persister phenotype following the activation of chromosomal *mazF* transcription, as well as after activation of *mazE-mazF* by antibiotic pre-treatment[248].

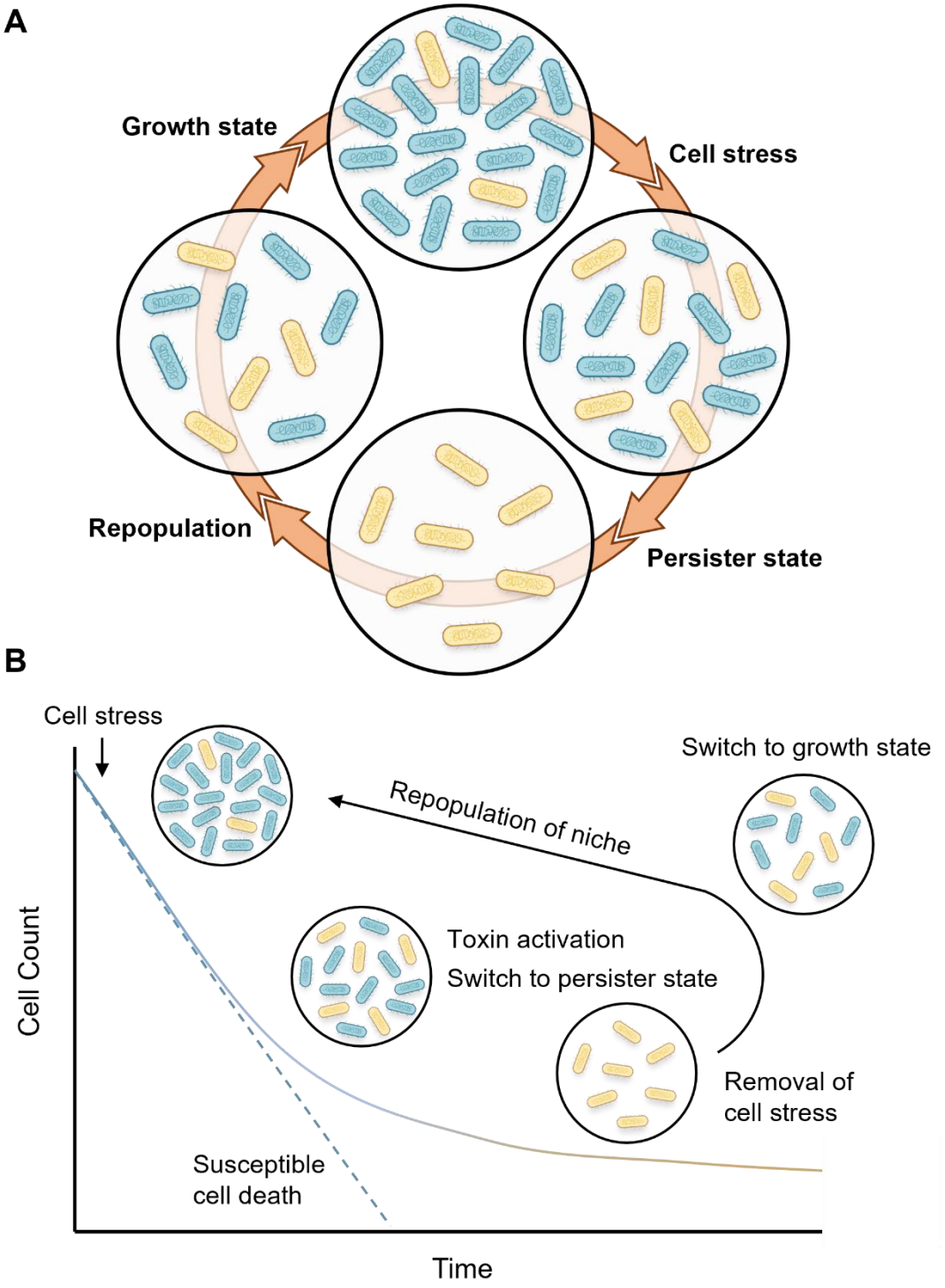


Figure 1.7. Schematic of the TA-mediated bacterial persistence model. (A) Growth state: actively growing cells (teal), with small, stochastic subpopulations of persister cells (yellow). Cell stress: upon cellular stress, i.e., nutrient starvation or drug exposure, unknown mechanisms contribute to toxin activation, leading to growth inhibition and a switch to the persister state. Persister state: as a consequence of cell stress, sensitive cells are killed, while highly tolerant persister cells survive. Repopulation: when cellular stress is removed, the antitoxin concentration increases, leading to toxin neutralisation, reversion to a growth state, and repopulation of the niche. **(B)** Representative biphasic killing curve depicting the survivability in response to stress of a homogeneously susceptible strain (dashed line) compared to a mixed population able to phenotypically switch to the persister state (solid line).

TA systems have been shown to contribute to persistence in organisms other than *E. coli*, significantly in pathogenic bacteria such as *Salmonella*[132,133,249,250]. Helaine *et al*. in 2014 reported that murine macrophage internalisation of *S*. Typhimurium led to the Lon-dependent activation of 14 chromosomal type II TA systems and a switch to the persister state, triggered by acidification and nutritional starvation within the macrophage vacuole[249]. Three of these TA systems encode TacT toxins, which acetylate aminoacyl-tRNAs to inhibit translation and induce persistence[132,133]. Additionally, ectopic expression of the *S*. Typhimurium ShpA toxin increases persistence in a Lon-dependent manner[250], whilst deletion of the *shpB-shpA* TA locus reduces persister formation compared to WT *S*. Typhimurium in infected mice[249]. Furthermore, expression profiles from drug-tolerant *M. tuberculosis* persister cells show an upregulation of TA genes[251,252].

Two paradigm-setting reports by the Gerdes group in 2011 and 2013, since retracted, strongly implicated TA systems in modulating persistence[253–256]. They found that the gradual deletion of ten chromosomal mRNase-encoding type II TA systems in *E. coli* resulted in a similarly gradual reduction in persister formation, with the deletion of all mRNase TA loci (Δ10) resulting in a 100-fold reduction in persister formation compared to WT levels[255,256]. They also observed that overproduction of Lon protease led to increased persister formation in WT cells, but had no effect in cells lacking the mRNase TA loci, the antitoxins of which were proposed as Lon substrates[255,256]. In the later 2013 study, also retracted, they further showed that stochastic variation of cellular (p)ppGpp levels affects persister formation, with high (p)ppGpp levels linked to stalled growth and a drug-tolerant phenotype[253,254]. This led to a persister cell formation model based on a hierarchical signalling cascade: stochastically high (p)ppGpp levels lead to Lon activation; Lon degrades antitoxins; free cognate mRNase toxins target translation to inhibit cell growth synergistically, leading to the persister state[253–256].

Whilst widely accepted, the Gerdes model also received significant scepticism from other researchers in the field[10,237,238,257]. Ramisetty *et al*. raised several issues: that Lon is not the sole protease responsible for the degradation of all type II antitoxins, undermining a fundamental principle of the model; that transcriptional regulation of the YefM-YoeB TA system, one of the original ten TA deletions implicated in persister formation, is not (p)ppGpp-dependent; and that the persister phenotype associated with the Δ10 strain was in fact fitness loss, arising from polar effects on important downstream genes due to the TA gene deletion methodology[238]. Similarly in response to this model, Goormaghtigh *et al*. independently re-assessed the earlier Gerdes lab studies and were unable to replicate a link between the *E. coli* mRNase TA systems and persistence[257].

Furthermore, a 2017 bombshell study by the Gerdes group, in which they re-tested their previously reported TA-mediated persister phenotypes[253–256], showed that these results were in fact due to inconsistencies with experimental procedures, assay induction of cryptic prophages, and bacteriophage ϕ80 contamination[258]. They instead found no clear evidence for TA involvement in persister formation, leading to the retraction of their previous work[253,256,258]. Additionally in 2017, the Lewis group observed that persistence occurs independently of (p)ppGpp and TA system activation[185]. They instead reported that persister formation is ATP-dependent, where stochastic decreases in ATP levels increase drug tolerance, presumably due to bactericidal antibiotics targeting energy-dependent processes[185].

Needless to say, the mechanisms underpinning persister formation are contentious and still unclear, with analyses often reductive in suggesting a simple signalling pathway or common molecular mechanism[257,259]. For example, gene expression profiling of *E. coli* and *M. tuberculosis* persister cells reveals a substantial number of differentially expressed genes relating to a range of stress responses[225,245,260]. Indeed, whilst numerous studies have demonstrated that specific TA systems can be involved in modulating persistence, our understanding of the full scale and mechanistic detail of their involvement is still limited, and questions of TA redundancy also need addressing. The observation that persistence increases in line with the cumulative deletion of TA modules, suggesting that these systems are redundant, is perhaps overly simplistic[257]. The knowledge that chromosomal TA modules reside primarily in the accessory genome, coupled with the large structural and mechanistic variety among even familial TA systems, suggests that a common, shared function such as persistence is unlikely[257]. Clearly, the intricacies of this relationship require further work.

1.5. Transcriptional autoregulation of TA systems

Transcriptional autoregulation is a feature of many TA loci, particularly type II TA systems[7]. The antitoxins of type II systems typically contain two domains: a DNA-binding domain, able to directly bind one or multiple operators within their promoter, thus repressing transcription, and a toxin-specific domain which is often intrinsically disordered until it binds the cognate toxin[7,261]. In many cases, antitoxin-toxin complexes can serve a purpose beyond simply stockpiling neutralised toxin; they have also been shown to bind DNA and enhance repression, adding further complexity to the regulatory repertoire[7,261].

1.5.1. Principle of conditional cooperativity

The most commonly reported autoregulatory process among type II systems is conditional cooperativity (Figure 1.8). The underlying mechanism behind conditional cooperativity was first described in a 1998 study of the Phd-Doc type II TA system[262]. The authors observed that the PhD antitoxin bound cooperatively as dimers (Phd_2) to either one or both palindromic sites within the cognate promoter, leading to partial transcriptional repression. The Doc toxin was in turn found to bind Phd_2, which enhanced Phd_2-DNA binding and strengthened transcriptional repression[262]. Interestingly however, they also observed that high concentrations of Doc relative to Phd resulted in no improvement in Phd_2-DNA binding cooperativity, and in fact led to partial derepression of transcription[262]. The cooperative DNA binding of Doc-Phd_2 repressor complexes is therefore conditional on the toxin to antitoxin ratio (ergo, conditional cooperativity): at low T:A ratios, the toxin co-represses transcription, whilst at high T:A ratios, it functions to destabilise and derepress[262,263] (Figure 1.8). These findings were replicated by a 2010 study which confirmed the role of conditional cooperativity in the autoregulation of *phd-doc*, and which also revealed that this process is contingent on the intrinsic disorder of the Phd antitoxin[264].

1.5.2. Intrinsic disorder and conditional cooperativity

Intrinsically disordered proteins (IDPs) are typically found in eukaryotes, where they play important roles in cell signalling pathways; IDPs, and the intrinsically disordered regions (IDRs) of some otherwise structured proteins, are susceptible to various PTMs and can allow a single protein to interact with multiple diverse partners[265]. In contrast, bacterial antitoxins comprise some of the few examples of prokaryotic proteins which feature IDRs[266–272]. The prevalence of IDRs in antitoxins is proposed to play a vital role in TA autoregulation (for an extensive review, see Loris and Garcia-Pino, 2014)[7,273]. In the case of the Phd-Doc system, Doc binding to the intrinsically disordered C-terminal domain (CTD) of Phd allosterically stabilises the antitoxin's N-terminal DNA-binding domain, shifting it towards a structured conformation which enhances DNA binding[264]. The 2010 Phd-Doc study found that *phd-doc* autoregulation is dependent both on T:A ratio, as well as the allosteric interplay between antitoxin-toxin binding and the intrinsic disorder of the Phd CTD[264]. At low T:A ratios, Doc bridges two separate operator-bound Phd_2 (Phd_2-Doc-Phd_2) to cooperatively bind DNA, increasing the avidity of Phd_2 for its operator sequence[264]. Doc binding to the Phd CTD also serves to stabilise and enhance Phd-DNA binding affinity, tightening repression[264]. At

saturating toxin levels, binding of surplus Doc leads to the formation of rigid Doc-Phd_2-Doc complexes which abolish cooperative DNA binding due to steric hindrance, causing derepression[264]. Conditional cooperativity has also been observed for several other type II systems, including CcdA-CcdB, Kis-Kid, and RelB-RelE, yet the core molecular mechanisms exhibit key differences[263,268,274,275].

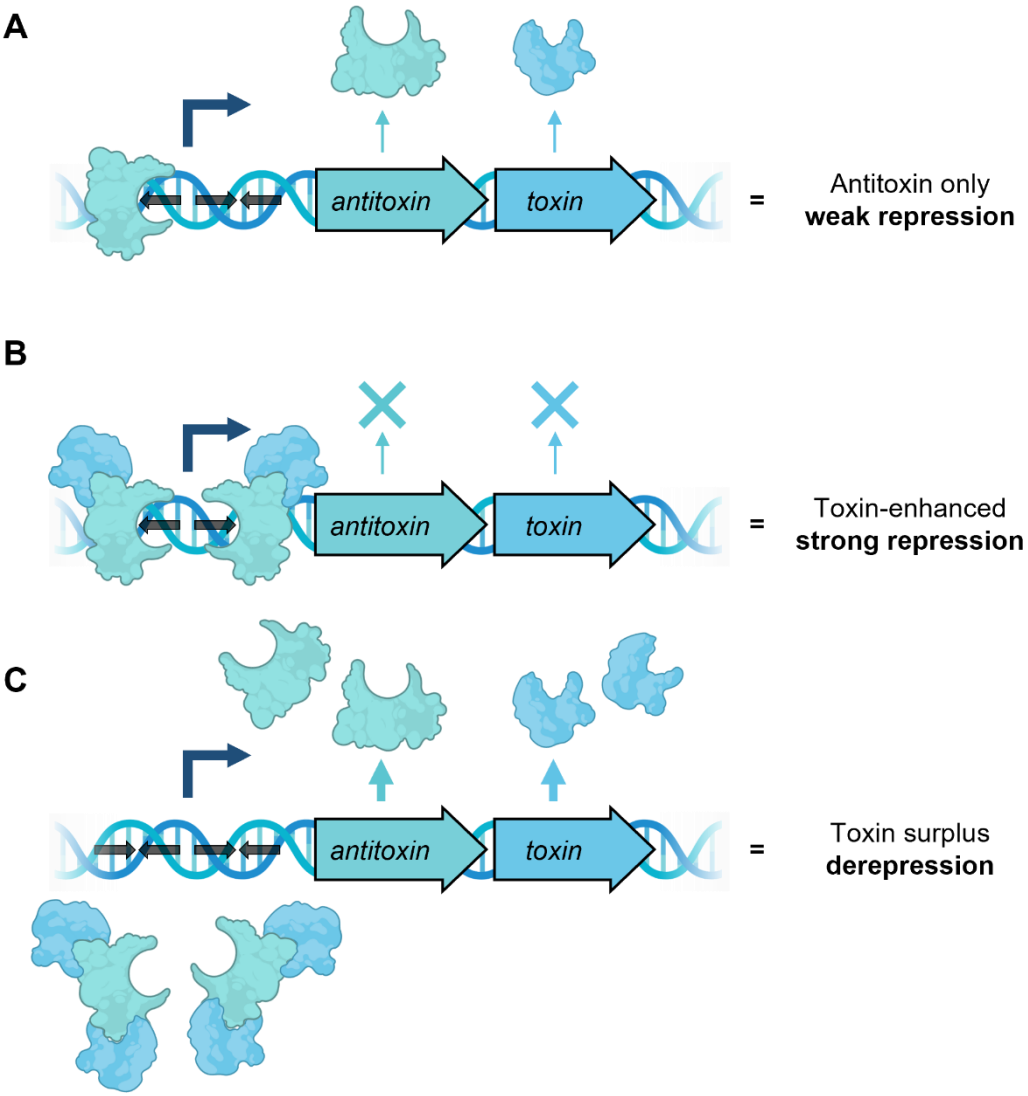


Figure 1.8. Schematic model of conditional cooperativity. (A) In the absence of the toxin, or at low T:A ratios, cognate antitoxins bind operator DNA in the TA promoter with low affinity, resulting in weak repression of transcription. **(B)** A similar T:A ratio leads to toxin-antitoxin binding, which enhances antitoxin-DNA binding affinity and tightens repression. **(C)** A surplus of toxin leads to alterations in the toxin-antitoxin complex architecture, which lowers antitoxin-DNA binding affinity and derepresses transcription.

1.5.3. Conditional cooperativity may provide advantageous adaptability

The potential benefits of conditional cooperativity to cell homeostasis are evident. Having transcriptional regulation linked to T:A ratio allows for automatic shifts that are reactive to environmental and cell context. Under normal growth conditions, the tight repression conferred by low T:A ratios helps to both manage resources and sequester the toxin to neutralise toxicity[65]. This mechanism can also function as an in-built feedback loop, protecting against fluctuating or accidental TA activation[65]. Most TA systems are expressed from a shared promoter, therefore any inappropriate increase in T:A ratio sufficient to cause derepression leads to both antitoxin and toxin transcription. Given that translation of antitoxins is typically preferential to toxin translation[276], derepression creates a feedback loop, whereby newly synthesised antitoxin can complex with available toxin to negate toxicity and repress transcription[65].

The inherent plasticity of the conditional cooperativity model also suggests a possible rationale for the involvement of TA systems in bacterial survival and stress response. For example, after sudden nutrient starvation – which is proposed to activate cellular proteases that preferentially degrade antitoxins – the T:A ratio may increase, leading to derepression and transcriptional activation of the TA operon[182,263]. Whilst under sustained stress, the newly expressed antitoxins are subject to proteolysis, thus preserving a high T:A ratio. This leaves toxin activity to continue unrestrained, both to inhibit growth and maintain derepression[263]. However, when stress abates and protease activation subsides, derepression leads to sustained antitoxin production, toxin neutralisation, and a return to a transcriptionally repressed state[263]. This model has also been offered as a possible explanation for the stochastic induction of the persister state in bacterial subpopulations[7,277], though this role has proved divisive (See, *1.4*). Conditional cooperativity could also represent an advantageous model for signal transmission; changes in protein ratios that alter the TA equilibrium could trigger a signalling cascade affecting numerous targets[264].

1.5.4. Alternative autoregulation

Not all type II TA modules are governed by conditional cooperativity; for example the DinJ-YafQ and MqsA-MqsR systems[278,279]. For the former, both dimeric DinJ ($DinJ_2$) and a tetrameric TA complex (YafQ-$DinJ_2$-YafQ) interact with an inverted repeat (IR) in the TA promoter, repressing transcription to the same extent[278]. In the case of MqsA-MqsR, and in contrast to previous reports indicating that the toxin functions as a co-repressor[89,153], MqsR

instead destabilises the MqsA-DNA complex to derepress transcription[279]. Interestingly, MqsR can cleave both *mqsA* and *mqsR* mRNA, highlighting a potential post-transcriptional mechanism of autoregulation beyond co-repression[279]. Indeed, other mechanisms that contribute to TA regulation and promote preferential antitoxin production abound at the post-transcriptional and translational level[14]. While these will not be discussed in great detail here, example mechanisms include transcriptional and translational coupling, operon architecture – such as number of promoters, location and relative strength of the toxin RBS, and presence of transcription terminators – and transcript number and completeness (i.e., full/unprocessed or truncated/active)[14]. Additionally, certain TA systems can themselves contribute to post-transcriptional and post-translational regulation: type I antitoxin sRNAs directly inhibit cognate toxin mRNAs and can promote their cleavage by RNases[280], whilst type VII antitoxins neutralise cognate toxins via chemical modifications[96,97,99].

1.6. Toxin activation

The exact conditions, mechanisms, and triggers that lead to TA system activation are somewhat unknown or else contested. Previous reports have highlighted proteolysis as a key component of TA activation. Proteolysis is a vital tool used by bacteria to maintain cellular homeostasis, as well as survive hostile environments. The onset of various stresses activates cellular housekeeping proteases – such as ATP-dependent Lon, ClpAP or ClpXP – which can dynamically reshape the proteome to promote adaptation to and recovery from stress[281–285]. A number of antitoxins have been shown to be substrates of these proteases, either by *in vitro* experiments demonstrating direct degradation or by displaying greater stability/longer half-lives when faced with a null mutant protease (Table 1.1). Previous TA activation models therefore proposed that when faced with certain stress conditions and stimuli, the activation of cellular proteases leads to the degradation of labile antitoxins. This frees the toxin to target vital cellular processes, thereby stalling growth and reducing metabolic burden until stress is relieved[7,162,182,286]. Various systems have been implicated under this model of activation via the observation that they are transcriptionally upregulated under conditions of stress, or when exposed to hostile environments such as those faced inside macrophages[12,167,168,181,182,249,287,288]. Notably, the upregulation of some TA modules during AA starvation has been shown to be Lon-dependent, lending support to this model[182,287].

Table 1.1. Antitoxins and the proteases implicated in their degradation

TA family	Antitoxin	Protease	Genetic origin	References
axe-txe	Axe1, Axe2	ClpCP	*S. aureus*	[289]
ccdA-ccdB	CcdA	Lon	F plasmid *E. coli*	[24,25]
dinJ-yafQ	DinJ	Lon, ClpXP	*E. coli*	[170]
ε-ζ	Epsilon	Lon, ClpXP*	pSM19035 *S. pyogenes*	[290]
hicB-hicA	HicB	Lon	*E. coli*	[183]
hipB-hipA	HipB	Lon, ClpP*, HslVU*	*E. coli*	[291]
mazE-mazF	MazE MazEsa	Lon, ClpAP ClpCP	*E. coli* *S. aureus*	[287,292] [289]
mqsA-mqsR	MqsA	Lon	*E. coli*	[152]
parD-parE	ParD	Lon	RK2 plasmid *E. coli*	[108]
pasABC	PasA	Lon	*E. coli*	[293]
pemI/kis-pemK/kid	PemI Kis	Lon ClpAP	R100 plasmid *E. coli* R1 plasmid *E. coli*	[294] [295]
phd-doc	PhD	ClpXP	P1 plasmid *E. coli*	[297]
relB-relE	RelB $RelB_{P307}$ RelN	Lon Lon Lon_s, $ClpP2_s/X_s$	*E. coli* P307 plasmid *E. coli* *Synechocystis* sp. PCC 6803	[182] [298] [299]
rnlB-rnlA	RnlB	Lon, ClpXP	*pTF-FC2 E. coli*	[300]
vapB-vapC	VapB	Lon	*S.* Typhimurium LT2	[301]
yefM-yoeB	YefM	Lon	*E. coli*	[302]

*Represents minor proteolytic effect. Table adapted from Brzozowska and Zielenkiewicz, 2013[296]

However, the common interpretation that stress-mediated upregulation of TA modules correlates to activation of the toxin component is proposed to be fundamentally flawed. Fraikin *et al*. noted that for many TA modules where expression has been found upregulated under stress conditions, no discernible growth arrest phenotype is in fact seen[57,171,182,198,200,287]. They highlighted that antitoxin translation has been found to be more efficient than that of the cognate toxin, therefore any derepression and upregulation of TA

modules should in theory produce an antitoxin surplus[14,198,276,303]. Hence, stress-induced upregulation of TA systems does not ensure an increased T:A ratio, and similarly does not guarantee toxin activity[198].

Certain groups have also queried whether antitoxins in a complexed state can be subjected to proteolysis, or how antitoxins might first disassociate from an antitoxin-toxin complex to then be degraded[20,198]. The toxin-binding domain of antitoxins is often intrinsically disordered, and this unstructured domain has previously been pinpointed as the target of protease recognition[7,9]. However, toxin binding allosterically shifts the antitoxin to a more stable conformation, making it unlikely that complexed antitoxins are preferentially degraded[20]. Indeed, several studies have shown that type II toxins and antitoxins form tightly bound compelxes[25,304,305]. These high affinity interactions can prevent access to cellular proteases, rendering the complexed antitoxin resistant to proteolysis[25,295,304].

A 2020 study by LeRoux *et al*. reaffirmed these points[21]. They demonstrated that ten chromosomal type II TA modules can be transcriptionally induced by a range of abiotic stress conditions, including AA starvation (SHX treatment), translation inhibition (chloramphenicol), oxidative stress (hydrogen peroxide), heat shock (30°C to 45°C), proteotoxicity (Δ*dnaK*), acid shock (pH 4), and inhibition of DNA synthesis (trimethoprim)[21]. Three of these systems were transcriptionally activated primarily via Lon-mediated antitoxin degradation, leading to increased T:A ratios, derepression and relief from autoregulation[21]. However, they also showed that these systems do not contribute to an altered growth phenotype under stress, nor that the RNase toxins significantly enhanced cellular RNA cleavage[21]. Complexed antitoxin was protected from proteolysis, with free antitoxin preferentially degraded[21]. These refractory pools of complexed antitoxin were proposed to keep the toxin sequestered and thereby prevent its liberation and activation; consistent with the lack of an appreciable toxin phenotype[21].

Considering the above, questions surrounding toxin activation linger. Given the ubiquity and diversity of TA systems in bacterial chromosomes, and the range of proposed physiological functions, a universal activating pathway induced by stress is unlikely. Increasingly, recent works have disputed this model, even in the context of previously tested systems. It is possible that additional and specific signalling may be required to activate distinct toxins, or promote the disassociation of TA complexes[21]. Environmental context is clearly important; TA systems have been shown to promote survival of *S*. Typhimurium in macrophages, therefore an unidentified biotic signalling determinant may be at large[21,249]. It has also been

proposed that *de novo* toxin synthesis is the primary cause of toxin activation and activity, in a background where antitoxin levels are insufficient for effective neutralisation, possibly due to post-translational inhibition of antitoxin mRNA[20].

Finally, the lack of a phenotype arising from toxin upregulation does not necessarily rule out toxin activity under physiologically relevant conditions. The growth inhibition phenotypes characteristic of toxin activity are usually the result of artificial, ectopic toxin expression, often plasmid based. In light of this, the effects of physiological levels of activated toxin may be less pronounced and may instead serve a more modest purpose. For example, exposing *E. coli* K-12 MC1000 (containing a chromosomal copy of *relB-relE*) to SHX treatment results in *relB-relE* transcriptional upregulation and a reduction in global translation rate without affecting cell viability[182]. Rather than completely shutting the cell down, this perhaps reflects a role reprogramming cell machinery or reallocating resources[287,306].

1.7. *M. tuberculosis* and TA systems

Various mentions in historical records describe tuberculosis (TB) under different names, with reports in Ancient Greek and Roman literature suggesting a hereditary cause and noting its contagious nature[307,308]. However, it was in 1882 that Robert Koch first identified the rod-shaped bacillus, *M. tuberculosis*, as the cause of TB[309]. Since Koch's discovery, palaeomicrobiology has revealed traces of *M. tuberculosis* in Neolithic human samples thousands of years old[310,311], providing a historical context for *M. tuberculosis* as an enduring human pathogen.

TB is a bacterial communicable disease which, in 2019, ranked among the top ten causes of death worldwide[312]. Primarily caused by *M. tuberculosis*, TB was notably the leading cause of death from a single causal agent and the leading infectious disease killer[312,313]. Indeed, in 2019, ten million people were estimated to have developed TB, and a further 1.4 million died [312], highlighting the prevalence of TB globally and the scale of the crisis.

1.7.1. *M. tuberculosis* pathogenicity and global burden

M. tuberculosis is typically transmitted human-to-human via aerosol droplets, where it is inhaled through the nose or mouth (Figure 1.9). The bacteria enter the lungs, where they are engulfed by alveolar macrophages and sequestered in phagosomes[314,315]. This primary infection is perhaps most characteristic of *M. tuberculosis* as a pathogen, where it forges a

commensal-like relationship with the host to survive. At this stage, *M*. *tuberculosis* must resist or control host immune responses, persist in the host whilst maintaining a relatively inoffensive infection, all while retaining the ability to reactivate and recommence infection at a later date[316].

Inside the macrophage, *M*. *tuberculosis* resides in a membrane-bound vacuole where it prevents maturation and acidification of the phagosome, and can even translocate to the cytosol[317–319]. In turn, granulomas form around the infected macrophages; these consist of an immune cell complex, comprising dendritic cells, granulocytes, natural killer cells, and T and B lymphocytes, which layer around the internal core of infected macrophages[320]. During sequestration in granulomas, *M*. *tuberculosis* presents as metabolically active but with a dormant growth state, persisting asymptomatically as part of the latent form of the disease[320]. To achieve this, *M*. *tuberculosis* carries an expansive regulatory network which allows it to respond to external stimuli and promote adaptation to the hostile environment within granulomas[321–326]. For example, *M*. *tuberculosis* features a range of secretion systems able to confuse and counter host immunity[327], whilst its distinctive cell wall – containing diverse lipids and glycolipids, as well as the characteristic mycolic acids – offers extreme hydrophobicity, as well as the ability to interact with host biological processes to potentially promote pathogenesis[316].

As a result, the organism is rarely eradicated. While *M. tuberculosis* is characterised by its ability to survive early immune responses and establish latent infection, another key characteristic is the organism's ability to later reactivate. Following a period of latency, *M. tuberculosis* can re-commence growth and replication at the initial site of infection, transitioning to the active state of TB (active TB) before disseminating to other parts of the body (extrapulmonary TB)[320]. This primarily occurs in immunocompromised individuals, where reactivated *M*. *tuberculosis* causes necrosis of infected macrophages, triggering fragmentation of the granuloma and release of actively replicating *M*. *tuberculosis*[320].

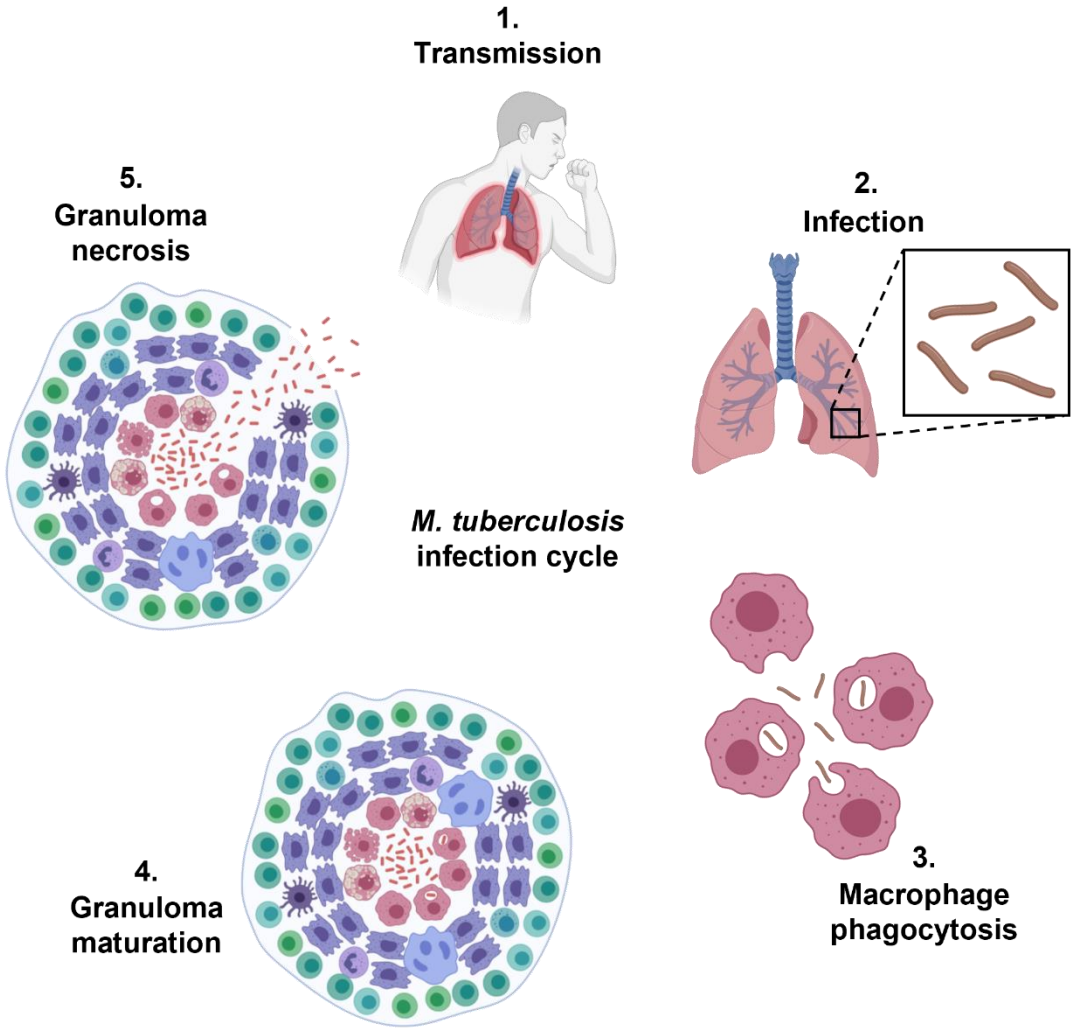


Figure 1.9. Simplified schematic of the *M. tuberculosis* infection cycle. (1) Transmission via aerosol droplets leads to infection of a new host's lungs. (2) Bacteria reproduce and infection is disseminated in lung tissue, prompting initiation of the immune response. (3) Immune cells are engaged, among them macrophages phagocytosing *M. tuberculosis* cells. The infection is either eradicated, or the highly tolerant *M. tuberculosis* bacteria are able to survive immune cell-mediated killing. (4) A granuloma is formed; an aggregate of immune cells surrounding dormant *M. tuberculosis* bacteria. This stage of the disease is known as latent TB, where the bacteria are contained and dormant. (5) In cases of reactivated TB, the necrotic granuloma ruptures, releasing actively replicating *M. tuberculosis* cells. Bacteria are disseminated and the infection is active and transmissible.

In 2014 and 2015, the World Health Organisation's (WHO) End TB Strategy and the United Nations' Sustainable Development Goals were implemented, which committed member nations to ending the TB epidemic by 2035 and 2030, respectively[328,329]. In recent years, improved awareness and international cooperation have resulted in a steady decrease in the number of TB deaths each year, with overall funding, reporting capabilities and early TB diagnoses generally increasing[312]. However, despite these and more recent international targets, in 2019, a quarter of the world's population was estimated to be infected with *M. tuberculosis*[312]. Among those infected with *M. tuberculosis*, estimates suggest between 5-10% stand the risk of developing active TB disease during their lifetime[312]. Perhaps most significantly, according to the WHO 87% of those who fell ill with active TB in 2019 were from 30 high TB burden countries[312]. Most of these are characterised as low- and middle-income countries, highlighting TB as a disease of the poor.

1.7.2. Drug resistance and development shortfalls

TB is both preventable and curable, however successful treatment relies upon access to speedy diagnosis, effective antimicrobials, and strict adherence to the treatment regimen. The latter two are especially important given the increasing prevalence of rifampicin-resistant TB (RR-TB) and multi-drug-resistant TB (MDR-TB; defined as rifampicin- and isoniazid-resistant), which renders first-line drugs ineffective[312]. The rise in drug-resistant *M. tuberculosis* infections can be linked to the complicated nature of first-line treatment[330]. This usually comprises a six month combination regimen of four first-line drugs: isoniazid, rifampicin, pyrazinamide and ethambutol[312]. The failure of patients to adhere to this regimen, due in part to duration, cost and drug toxicity, results in selection for resistant *M. tuberculosis* variants which are considerably more problematic[330]. In 2019 almost 500,000 people developed RR-TB, and of these cases 78% were classified as MDR-TB[312]. The global treatment success rate of RR-/MDR-TB is only approximately 57%, and treatment itself is substantially more complicated; the required medication regime is longer, lasting between nine and twenty months, and the drugs used are both more costly (≥ US$ 1000 per/person) and more toxic to the recipient[312]. The rise of extensively drug-resistant TB (XDR-TB) further complicates treatment. The recently updated definition of XDR-TB describes TB cases caused by *M. tuberculosis* strains which are MDR, and also carry resistance to second-line drugs including fluoroquinolones[331].

Despite the evident need, there is a glaring shortage of *M*. *tuberculosis* drugs and vaccines in the developmental pipelines of major pharmaceutical companies. According to WHO figures, 22 drug and 14 vaccine candidates were undergoing clinical trials worldwide as of August 2020, which initially seems promising[312]. However, as of March 2021, a search of drug development pipelines of six major Western pharmaceutical companies revealed that only 3 out of 730 products (0.4%) were related to TB (Figure 1.10). More significantly, the only currently licensed vaccine remains the 100-year-old bacilli Calmette-Guérin (BCG) vaccine, used for preventative TB treatment in children[312]. No vaccine exhibiting efficacy in adults has so far been produced, whilst the last TB drugs approved by the US Food and Drug Administration were the rifamycins, in 1971[330].

1.7.3. An abundance of TA systems: innate influencers of latent infection?

Besides the prevalence of drug-resistant TB, one of the primary factors driving recalcitrant *M*. *tuberculosis* infections is its characteristic ability to transition to a dormant growth state, surviving inside macrophages, before later reactivating to an active growth state[316]. This dormant persister-like state is refractory to immune responses and antibiotic killing, and thought to contribute to the formation of latent TB[11].

M. *tuberculosis* is distinct among bacteria in boasting the greatest number of chromosomal TA systems, with upwards of eighty identified so far, predominantly encoded on genomic islands[11,12] (Figure 1.11). When activated, TA systems typically exert a reversible, bacteriostatic effect on bacterial growth. They have been linked to persister formation in many organisms (See, *1.4*), as well as bacterial stress responses (See, *1.3.3*), and mediating the intracellular survival of pathogenic *S*. Typhimurium inside macrophages[249]. These pathogenic traits are shared by *M*. *tuberculosis*, which has therefore provoked interest in the role of TA systems in these phenotypes[11,12]. Notably, many of the TA modules encoded by *M*. *tuberculosis* are conserved only in the *M*. *tuberculosis* complex (MTBC), a genetically related group of pathogenic mycobacteria, suggesting a potentially important role in *M*. *tuberculosis* evolution and pathogenesis[12]. In contrast, the genome of *M. smegmatis*, a non-pathogenic bacterium of the soil, harbours just five TA modules[13].

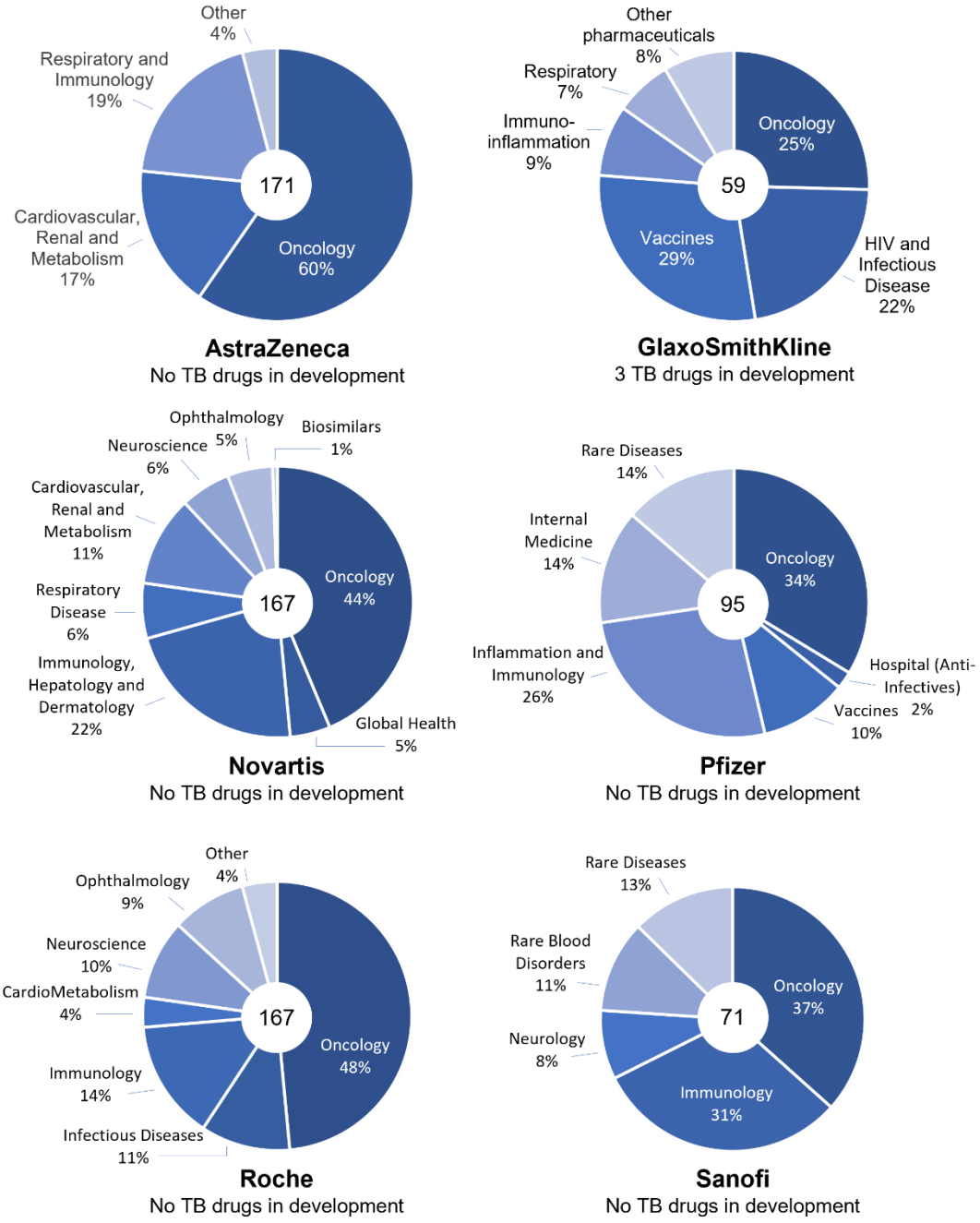


Figure 1.10. Drug development pipelines of six major pharmaceutical companies grouped into research area. Developmental drugs are grouped into research areas as described by each individual company and depicted as a percentage of the total number of products (central values). Of the companies selected, a total of 730 products were in development pipelines at the time of analysis, with three of these targeting TB (0.4%). Data are publicly available and were accessed on, and are accurate as of, 09/03/2021.

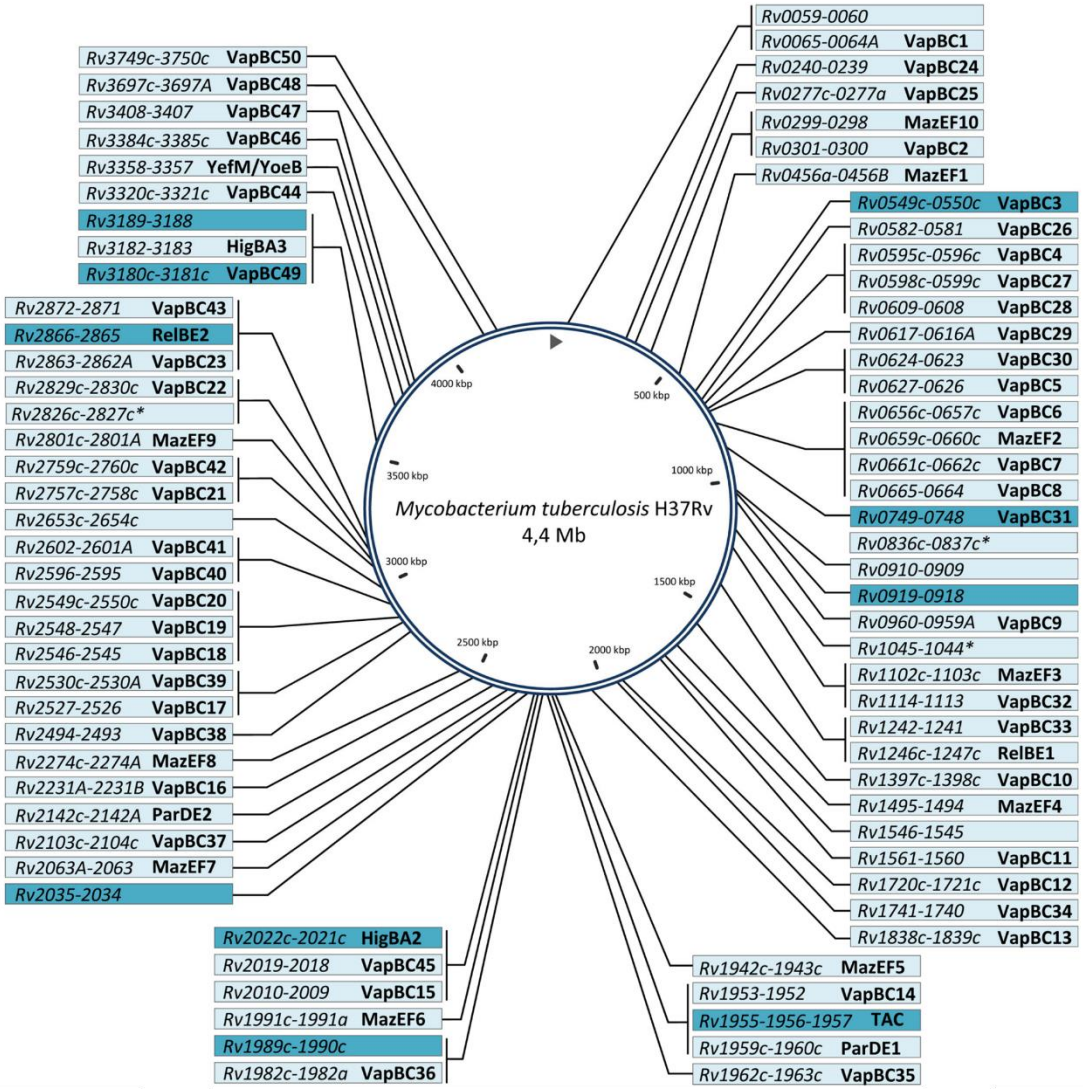


Figure 1.11. *M. tuberculosis* H37Rv chromosomally encoded TA systems. The TA systems of *M. tuberculosis* are predominantly type II; notable exceptions are marked by an asterisk (*) and are proposed to be putative type IV modules. Systems imposed on a dark blue background represent the ten most induced TA modules in *M. tuberculosis* persister cells. Figure adapted from Sala *et al*., 2014[11].

M. *tuberculosis* TA systems are predominantly type II RNase-encoding modules; amongst these, the VapB-VapC super-family are the most abundant[11] (Figure 1.11). The sheer number of familial *M*. *tuberculosis* type II systems raises questions as to their role: they may function in genome stabilisation, as discussed previously, or alternatively individual TA systems may be regulated via independent mechanisms to contribute to pathogencity[12]. Different *M*. *tuberculosis* RNase toxins may have unique cleavage sites, allowing condition-specific degradation of RNAs that accumulate under explicit stresses; in turn, toxin-mediated

transcriptome remodelling may contribute to intracellular messaging, allowing subtle fine-tuning of metabolic status, promoting phenotypic heterogeneity, and providing adaptability to host challenge[12,332].

Indeed, different *M*. *tuberculosis* TA systems are activated in response to diverse host-associated environmental stresses. Expression profiling has demonstrated that several *M*. *tuberculosis* TA systems are upregulated during hypoxia, which replicates the intra-macrophage environment, as well as during macrophage infection[12]. Interestingly, the TA systems induced under each of these conditions do not overlap, indicating that the regulation of TA modules is independent of one another, activated by different environmental triggers within the macrophage[12]. Additionally, both the SR and SOS response have been reported to activate *M*. *tuberculosis* TA systems[332,333]. Transposon site hybridisation (TraSH) studies have also highlighted TA systems that are required for slow growth, whilst a number of TA systems have been found upregulated in drug-tolerant *M*. *tuberculosis* persister cells[11,251,252,260]. Other TA modules are reportedly important for regulating intracellular growth and replication, secretion of immunogenic proteins, virulence, as well as promoting intra-macrophage survival (for comprehensive reviews, see Sala *et al*., 2014, and Slayden *et al*., 2018)[11,332]. As a result, enhancing our understanding of how *M*. *tuberculosis* TA systems contribute to bacterial physiology, and identifying new members of its broad TA repertoire, may provide clues to better control and treat this problematic pathogen.

1.8. Research aims

M. *tuberculosis* presents a continuing, major challenge to global health. Enhancing our understanding of how intrinsic bacterial systems contribute to the almost unrivalled survivability of *M*. *tuberculosis* is therefore vital, and may help to develop and refine strategies that target chronic *M*. *tuberculosis* infections.

This project originally focused on a family of three putative TA systems from *M*. *tuberculosis*: Rv0837c-Rv0836c, Rv1044-Rv1045, and Rv2827c-Rv2826c, herein described as $MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$, respectively. These systems were originally identified in 2014 by Dy *et al*., where $MenT_3$ and $MenT_4$ exhibited homology to the type IV AbiEii toxin from *S. agalactiae* in the form of a conserved NTase-like DUF1814 domain[83]. A second 2014 study by Sala *et al*. also identified $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ among the broad repertoire of *M*. *tuberculosis* TA systems, and further highlighted $MenT_2$ as a DUF1814 protein[11]. Notably, these three systems represented the first examples of putative type IV TA modules in *M*. *tuberculosis*[11].

These systems are of particular interest due to their reported involvement in physiologically relevant phenotypes. Several transcriptome profiling experiments have shown that *$menA_2$*, *$menT_2$*, *$menA_3$*, and *$menA_4$* are upregulated in *M*. *tuberculosis* persister populations, though not consistently across all studies[251,252,260]. Furthermore, a previous TraSH study found that $MenA_4$ is essential for *M*. *tuberculosis* growth[334], whilst in 2017, saturating transposon mutagenesis experiments revealed that insertions in *$menA_4$* cause a growth defect, and that *$menA_3$* is also essential[335]. The observation that both $MenA_3$ and $MenA_4$ antitoxins are necessary for growth strongly suggests that the putative cognate toxins, $MenT_3$ and $MenT_4$, are active and toxic in *M*. *tuberculosis*.

At the project's outset, the mode of activity and role of these three putative TA modules remained uncharacterised[11]. Homologous systems from different bacteria had previously been tested and demonstrated typical TA activity, defined by reversible toxicity[83,336]. As a result, the first aim of this project was to functionally characterise $MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$, and determine if they do indeed function as TA systems.

AbiEi-AbiEii is reportedly a type IV TA system, in which the antitoxin and toxin do not interact but instead work antagonistically to neutralise and effect toxicity, respectively[83]. Interestingly, AbiEi can also bind the *abiEi-abiEii* promoter to repress transcription and autoregulate its own operon, similar to type II systems[83,84]. Additional project aims therefore

sought to establish if antitoxin-mediated autoregulation is similarly a feature of the MenA-MenT systems, and to define the TA class to which the *M*. *tuberculosis* homologues belong.

The AbiEii toxin acts as a GTP-specific NTase to exert toxicity, however its target remains elusive[83]. As a result, a major project aim involved determining the cellular target of the homologous $MenT_2$, $MenT_3$, and $MenT_4$ toxins, and elucidate their mode of toxicity. In accordance with these aims, this project utilised a range of microbiological, biochemical, and structural techniques to explore and characterise these systems.

Finally, a fourth MenA-MenT family TA system – annotated as Rv0078B-Rv0078A and renamed $MenA_1$-$MenT_1$ – was later identified and functionally characterised by collaborators during the course of this work. Consequently, an additional aim of this project was to structurally characterise $MenA_1$-$MenT_1$ and determine the extent to which core MenA-MenT features are conserved within this family.

2. Materials and Methods

2.1. Media, solutions, and chemicals

The media and solutions used in this study are described in Tables 2.1 and 2.2, respectively. The antibiotics and supplements used, alongside stock and working concentrations, are listed in Table 2.3. Where appropriate, media and solutions were sterilised by autoclaving at 121 °C for 15 min, or filter sterilised using 0.22 µM filters (Fisher Scientific). For visualisation and analysis by SDS-PAGE, staining was performed with either InstantBlue Protein Stain (Expedeon; Cat No. ISB1L) or Quick Coomassie Stain (Neo Biotech).

2.2. Bacterial strains and culture conditions

The bacterial strains used in this study are listed in Table 2.4. *E. coli* DH5α were used for cloning and functional assays; *E. coli* ER2566 and BL21(λDE3)Δ*slyD* were used to express recombinant proteins for purification, biochemistry, and crystallisation. *E. coli* strains were typically grown at 37 °C, either on LBA or M9A solid media or in LB, M9M or 2x YT liquid culture (Table 2.1), with shaking at 180 rpm. The growth of liquid cultures was measured by optical density using a Biochrom WPA CO8000 spectrophotometer set to 600 nm wavelength (OD_{600}). Where necessary, bacteria were supplemented with antibiotics and inducing or repressing agents (Table 2.3). To prepare bacterial strains for long term storage, 500 µl of overnight culture was mixed with 500 µl 50% glycerol in a 1.8 ml Cryovial (Starlab) and snap-frozen in liquid N_2, then relocated to a -80 °C freezer.

2.3. DNA manipulation

All DNA was isolated and manipulated using standard molecular biology techniques. Unless otherwise stated, all kits were used according to the manufacturers' instructions. Oligonucleotides were purchased from Integrated DNA Technologies (IDT) and are listed in Table 2.5. *M. tuberculosis* H37Rv genomic DNA was obtained from ATCC. All plasmids used in this study are detailed in Table 2.6. Where the plasmid source was external to this project, the construction notes are as per the given reference. Plasmids generated as a result of this study were constructed using a mixture of the techniques and kits described below, and are described in greater detail within the appropriate Results section corresponding to their first use.

Table 2.1. Media used in this study

Medium	Ingredients per litre (dH_2O)
Luria broth (LB) (Melford)	10 g casein digest peptone 10 g NaCl 5 g yeast extract Autoclave sterilised
LB agar (LBA) (Melford)	10 g casein digest peptone 10 g NaCl 5 g yeast extract 15 g agar Autoclave sterilised
2x YT broth (Melford)	16 g digest peptone 10 g yeast extract 5 g NaCl Autoclave sterilised
Super optimal broth (SOB) (Melford)	20 g casein enzymic hydrolysate 2.4 g $MgSO_4$ 5 g yeast extract 0.186 g KCl 0.5 g NaCl Autoclave sterilised
5x M9 minimal salts	33.9 g Na_2HPO_4 15 g KH_2PO_4 2.5 g NaCl 5 g NH_4Cl
Supplemented M9 minimal media (M9M)	200 mL 5x M9 minimal salts 20 mL 20% D-glu * 1 mL 0.05 % thiamine HCl 1 mL 15 mM $FeSO_4$ 1 mL 1 M $MgSO_4$ Autoclave sterilised **Note**: addition of 6 g agar if making M9 minimal agar (M9A) * where D-glu was required for repression of P_{BAD}-mediated expression, 1% glycerol was used

Table 2.2. Solutions used in this study

Solution	Ingredients
All solutions were made using Milli-Q water (MQ) unless otherwise stated	
DNA work	
Solution A (chemically competent cells)	9.9 mM $MnCl_2$ 49.5 mM $CaCl_2$ 9.9 mM MES Autoclave sterilised
Solution B (chemically competent cells)	9.9 mM $MnCl_2$ 49.5 mM $CaCl_2$ 9.9 mM MES 50% glycerol Autoclave sterilised
Inoue buffer	55 mM $MnCl_2$ 15 mM $CaCl_2$ 250 mM KCl 10 mM PIPES pH 6.7 Filter sterilised
50x TAE buffer (Severn Biotech Ltd)	1x working concentration: 40 mM tris acetate 1 mM EDTA pH 8.0
DNA Gel Loading Dye (6x) (Thermo Fisher Scientific)	N/A, used as per manufacturer's instructions
Agarose gel mix	1% agarose (Severn Biotech Ltd) dissolved in TAE, followed by addition of 500 ng/ml ethidium bromide
Protein purification and crystal production	
A500 buffer	20 mM Tris HCl pH 7.9 500 mM NaCl 5 mM imidazole pH 8.0 10% glycerol
A100 buffer	20 mM Tris HCl pH 7.9 100 mM NaCl 5 mM imidazole pH 8.0 10% glycerol
B100 buffer	20 mM Tris HCl pH 7.9 100 mM NaCl 250 mM imidazole pH 8.0 10% glycerol
C1000 buffer	20 mM Tris HCl pH 7.9 1 M NaCl 10% glycerol
Sizing buffer	50 mM Tris HCl pH 7.9 500 mM KCl 10% glycerol

Table 2.2 continued. Solutions used in this study

Solution	Ingredients
All solutions were made using MQ unless otherwise stated	
Protein purification and crystal production	
Storage buffer	50 mM Tris HCl pH 7.9 500 mM KCl 70% glycerol
Crystal buffer	20 mM Tris HCl pH 7.9 150 mM NaCl 2.5 mM DTT
Cryo buffer	25 mM Tris HCl pH 7.9 187.5 mM NaCl 3.125 mM DTT 80% glycerol
Selenomethionine protein production	
Nutrient mix	4 mg/ml L-lysine hydrate 4 mg/ml L-threonine 4 mg/ml L-phenylalanine 2 mg/ml L-leucine 2 mg/ml L-isoleucine 2 mg/ml L-valine 4 mM $CaCl_2$ Filter sterilised
SelenoMethionine Solution (250x) (Molecular Dimensions)	Added to growth medium to a working concentration of 40 µg/ml
One-dimensional SDS-PAGE	
10x SDS running buffer	248 mM Tris base 1.92 M glycine 35 mM SDS powder Note: for 1x, dilute 1:10 v/v in MQ
15% acrylamide gel (resolving) (enough for two 10 cm Bio-Rad gels)	5.7 ml dH_2O 6 ml 40% acrylamide (Severn Biotech Ltd) 4 ml 1.5 M Tris HCl pH 8.8 160 µl 10% SDS 160 µl 10% ammonium persulphate 16 µl TEMED
4% acrylamide gel (stacking) (enough for two 10 cm Bio-Rad gels)	6.3 ml dH_2O 1 ml 40% acrylamide (Severn Biotech Ltd) 2.5 ml 0.5 M Tris HCl pH 6.8 100 µl 10% SDS 100 µl 10% ammonium persulphate 10 µl TEMED

Table 2.2 continued. Solutions used in this study

Solution	Ingredients
All solutions were made using MQ unless otherwise stated	
One-dimensional SDS-PAGE	
4x SDS loading buffer	40 mM Tris HCl pH 6.8 40% glycerol 4 mM EDTA 2.5% SDS 0.2 mg/ml Bromophenol blue 125 mM DTT Stored at -20 °C
InstantBlue Protein Stain (Expedeon)	N/A, used as per manufacturer's instructions
Quick Coomassie Stain (Generon)	N/A, used as per manufacturer's instructions
Other	
Phosphate-buffered saline (PBS) (Melford)	Used as per manufacturer's instructions, dissolving 1 tablet per 100 mL dH_2O, yielding: 11.9 mM phosphate buffer (Na_2HPO_4 and NaH_2PO_4) 2.7 mM KCl 137 mM NaCl pH 7.4

Table 2.3. Antibiotics and supplements used in this study

Chemical (abbreviation)	Stock solution (prepared in dH_2O unless otherwise stated)	Working concentration
Antibiotic		
Ampicillin (Ap)	1000x: 100 mg/ml, stored at -20 °C	100 µg/ml
Chloramphenicol (Cm)	1000x: 25 mg/ml in 100% EtOH, stored at -20 °C	25 µg/ml
Kanamycin (Km)	1000x: 50 mg/ml, stored at -20 °C	50 µg/ml
Spectinomycin (Sp)	1000x: 100 mg/ml, stored at -20 °C	100 µg/ml
Tetracycline (Tc)	1000x: 10 mg/ml in 50% EtOH v/v, stored at -20 °C	10 µg/ml
Supplement		
Isopropyl β-D-thiogalactopyranoside (IPTG)	1000x: 1 M, stored at 20 °C	1 mM
L-arabinose (L-ara)	100x: 10% w/v, autoclave sterilised	0.1%
D-glucose (D-glu)	100x: 20% w/v, autoclave sterilised	0.2%

Table 2.4. Bacterial strains used in this study

Strain	Genotype/characteristics	Source
Escherichia coli		
BL21(λDE3)Δ*slyD*	B; F^- *ompT gal dcm lon* $hsdS_B(r_B^-m_B^-)$ λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) $[malB^+]_{K\text{-}12}(\lambda^S)$ Δ*slyD*::Km^R	Genevaux Lab
DH5α	K-12; F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *hsdR*17($r_K^-m_K^+$) λ^-	Invitrogen
ER2566	B; F^- λ^- *fhuA2* [*lon*] *ompT lacZ*::*T7p07 gal sulA11* Δ(*mcrC-mrr*)114::IS*10* R(*mcr*-73::miniTn*10*-Tc^S)2 R(*zgb-210*::Tn*10*)(Tc^S) *endA1* [*dcm*]	New England Biolabs

Table 2.5. Oligonucleotides used in this study

Name	5′ – 3′ Nucleotide sequence	Notes	Purpose
TRB1018	CAACAGCAGACGGGAGGTTGTGCAAAACCGTATCTAATTGATACGATTGCGC	Fwd LIC-*menA*$_3$	pTRB491 cloning
TRB1019	GCGAGAACCAAGGAAAGGTTATTACGCCGATGCTCGCTTCGG	Rev LIC-*menA*$_3$	pTRB491 cloning
TRB1020	CAACAGCAGACGGGAGGTACCAAGCCCTATTCGTCGCCG	Fwd LIC-*menT*$_3$	pTRB492, pTRB576 and pTRB577 cloning
TRB1021	GCGAGAACCAAGGAAAGGTTATTATCTTTTCGTCGCCCGATCAATCCG	Rev LIC-*menT*$_3$	pTRB492, pTRB576, pTRB577, pTRB599, and pTRB631 cloning
TRB1022	CAACAGCAGACGGGAGGTGTGAGCCCAGCCGGCGCC	Fwd LIC-*menA*$_4$	pTRB493 cloning
TRB1023	GCGAGAACCAAGGAAAGGTTATTACGCCTTGCCGATCACGCGCAGC	Rev LIC-*menA*$_4$	pTRB493 cloning
TRB1024	CAACAGCAGACGGGAGGTGCCGGTCTGACCCGTGCGCTC	Fwd LIC-*menT*$_4$	pTRB494 and pTRB575 cloning
TRB1025	GCGAGAACCAAGGAAAGGTTATTAGGACCGCAGCACCGCCAGC	Rev LIC-*menT*$_4$	pTRB494 and pTRB575 cloning
TRB1028	CAACAGCAGACGGGAGGTCTCGTCGGGGCACAGTGCCGC	Fwd LIC-*menT*$_2$	pTRB496 cloning
TRB1029	GCGAGAACCAAGGAAAGGTTATTAGCCCGGTCGACCCACGGCGAAG	Rev LIC-*menT*$_2$	pTRB496 cloning
TRB1030	TTTGAATTCAGGAGGACAGGGATGCTCGTCGGGGCACAGTGC	Fwd EcoRI-*menT*$_2$	pTRB480 cloning
TRB1031	TTTAAGCTTGGCTAGCCCGGTCGACCCACGG	Rev HindIII-*menT*$_2$	pTRB480 cloning
TRB1040	TTTGAATTCGGGTCCCAACCGAGCGGC	Fwd EcoRI-P*menA*$_3$	pTRB483 cloning
TRB1041	TTTAAGCTTATTAGGTGATGGAGGCCAAGGCC	Rev HindIII-P*menA*$_3$	pTRB483 cloning
TRB1042	TTTGAATTCGCCCAAGCATCGGCTGGC	Fwd EcoRI-P*menA*$_4$	pTRB484 cloning
TRB1043	TTTAAGCTTCCGAACTTGAATTCACACCGG	Rev HindIII-P*menA*$_4$	pTRB484 cloning
TRB1044	TTTGAATTCACCTCGTCGAGTGGGTCG	Fwd EcoRI-P*menA*$_2$	pTRB485 cloning
TRB1047	TTTAAGCTTTACGGCCCCCCACTTGTTGC	Rev HindIII-P*abiEi*	pTRB486 cloning
TRB1052	TTTGAATTCAGGAGGACAGGGATGTCAAAAAAAGAGATTCTACTC	Fwd EcoRI-*abiEi*	pTRB481 cloning
TRB1053	TTTAAGCTTGGTTATTATATTAGAACCTCCAGAGTTTG	Rev HindIII-*abiEi*	pTRB481 cloning

Table 2.5 continued. Oligonucleotides used in this study

Name	5′ – 3′ Nucleotide sequence	Notes	Purpose
TRB1054	TTTGAATTCAGGAGGACAGGGATGAACAAAG CTAAATTAACAGC	Fwd EcoRI-*abiEii*	pTRB482 cloning
TRB1055	TTTAAGCTTGGTTATTATTAATCAGAGTCAGTA GAAATAATAATTG	Rev HindIII-*abiEii*	pTRB482 cloning
TRB1072	TTTGAATTCGATTTTGTTATCACAATAAATTGA GG	Fwd EcoRI-P_{abiEi}	pTRB486 cloning
TRB1073	TTTGAATTCTGCGTTCAGTATATCCGTACGTTC AG	Rev EcoRI-P_{menA_2}	pTRB485 cloning
TRB1120	TTTGGTACCAAGAAGGAGATATATCCATGAGT GGC	Fwd KpnI-RBS-His_6-SUMO	pTRB517 and pTRB545 cloning
TRB1121	GCCTCCCGTCTGCTGTTGAA	Rev His_6-SUMO	pTRB517, pTRB544 and pTRB545 cloning
TRB1122	TTCAACAGCAGACGGGAGGCACCAAGCCCTA TTCGTCGCC	Fwd His_6-SUMO end, *menT*$_3$ start	pTRB517 cloning
TRB1124	TTTAAGCTTTTATTATCTTTTCGTCGCCCGATC AA	Rev HindIII-*menT*$_3$	pTRB517 cloning
TRB1175	TTTCCCGGGAAGAAGGAGATATATCCATGAGT GGC	Fwd SmaI-RBS-His_6-SUMO	pTRB544 cloning
TRB1176	TTCAACAGCAGACGGGAGGCGCCGGTCTGAC CCGTGCG	Fwd His_6-SUMO end, *menT*$_4$ start	pTRB544 cloning
TRB1177	TTTAAGCTTTTATTAGGACCGCAGCACCGCCA G	Rev HindIII-*menT*$_4$	pTRB544 cloning
TRB1321	TTTCCCGGGAAGAAGGAG	Fwd XmaI-*menT*$_4$	pTRB544 cloning
TRB1322	TTTAAGCTTTTATTAGGACCGC	Rev HindIII-*menT*$_4$	pTRB544 cloning
TRB1460	TTTGAATTCTTTGTTTAACTTTAAGAAGGAGAT ATATCC	Fwd EcoRI-RBS-His_6-SUMO-LIC site	pTRB548, pTRB550 and pTRB596 cloning
TRB1461	TTTGAATTCGATGAAGATACAATTGATGTGTT TCAAC	Fwd EcoRI-LIC site	pTRB546, pTRB547, pTRB549, pTRB551 and pTRB595 cloning
TRB1462	TTTAAGCTTGGATCCCTCGAGGTCGAC	Rev HindIII-LIC site	pTRB546, pTRB547, pTRB548, pTRB549, pTRB550, pTRB551, pTRB595 and pTRB596 cloning
TRB1519	GCTGTTGGTCGCAGGCGGATCGTCGCTG	Fwd SDM-*menT*$_3$ K61A	pTRB559 cloning

Table 2.5 continued. Oligonucleotides used in this study

Name	5′ – 3′ Nucleotide sequence	Notes	Purpose
TRB1520	AGCGGAGCGCCGGTGTCG	Rev SDM-*menT₃* K61A	pTRB559 cloning
TRB1521	CATTCCCTGGGCAATCGCGCAGAAG	Fwd SDM-*menT₃* Q185A	pTRB560 and pTRB571 cloning
TRB1522	GTCATGCAGGGTACGGCG	Rev SDM-*menT₃* Q185A	pTRB560 and pTRB571 cloning
TRB1523	GCAAATCGCGGCAAAGCTGCACGC	Fwd SDM-*menT₃* Q188A	pTRB561 and pTRB572 cloning
TRB1524	CAGGGAATGGTCATGCAG	Rev SDM-*menT₃* Q188A	pTRB561 and pTRB572 cloning
TRB1525	AATCGCGCAGGCACTGCACGCAG	Fwd SDM-*menT₃* K189A	pTRB562 cloning
TRB1526	TGCCAGGGAATGGTCATG	Rev SDM-*menT₃* K189A	pTRB562 cloning
TRB1674	GACCTGGTGGCATTGCAGCTTCTTGAAG	Fwd SDM-*menT₃* D211A	pTRB592 and pTRB594 cloning
TRB1675	GTGAGCGCGGTCGTTGAC	Rev SDM-*menT₃* D211A	pTRB592 and pTRB594 cloning
TRB1676	ACCTCCAAAGCATTCGACACGGTCG	Fwd SDM-*menT₃* D80A	pTRB591 and pTRB593 cloning
TRB1677	CCGCGAATCGGGAATTCC	Rev SDM-*menT₃* D80A	pTRB591 and pTRB593 cloning
TRB1678	CAACAGCAGACGGGAGGTAGGAGGACAGGGATGACCAAGCCCTATTCGTCG	Fwd LIC-RBS-*menT₃*	pTRB599 and pTRB631 cloning
TRB1681	CAACAGCAGACGGGAGGTAGGAGGACAGGGATGGCAGTTTCCGTCGCTG	Fwd LIC-RBS-*menA₁*	pTRB597 cloning
TRB1682	GCGAGAACCAAGGAAAGGTTATTATCATGTGAACCGTGTGGACG	Rev LIC-*menA₁*	pTRB597 cloning
TRB1699	CAACAGCAGACGGGAGGTGCAGTTTCCGTCGCTGCG	Fwd LIC-*menA₁*	pTRB617 cloning
TRB1700	GCGAGAACCAAGGAAAGGTTATTATGTGAACCGTGTGGACGTC	Rev LIC-*menA₁*	pTRB617 cloning
TRB1701	CAACAGCAGACGGGAGGTAACGCTGTGGAGTCGACAC	Fwd LIC-*menT₁*	pTRB629 cloning
TRB1702	GCGAGAACCAAGGAAAGGTTATTACCACTTGGCGGCGAGGCG	Rev LIC-*menT₁*	pTRB629 cloning
TRB1726	CGGGCGGTTCGCAACTGATCTGGACTTC	Fwd SDM-*menT₄* S67A	pTRB618 cloning
TRB1727	ACGTTGCCCAGCCGGCAC	Rev SDM-*menT₄* S67A	pTRB618 cloning
TRB1728	GTTCTCCACTGCACTGGACTTCAGCGCGC	Fwd SDM-*menT₄* D69A	pTRB619 cloning

Table 2.5 continued. Oligonucleotides used in this study

Name	5′ – 3′ Nucleotide sequence	Notes	Purpose
TRB1729	CGCCCGACGTTGCCCAGC	Rev SDM-*menT₄* D69A	pTRB619 cloning
TRB1730	CTGCGCAGAGGCACTCGCGCGGT	Fwd SDM-*menT₄* K171A	pTRB620 cloning
TRB1731	GCTTCCGCTTCTGCCACG	Rev SDM-*menT₄* K171A	pTRB620 cloning
TRB1732	GGACCTGTACGCACTGAACCACTTCGCCTCGC	Fwd SDM-*menT₄* D186A	pTRB621 cloning
TRB1733	CGTGCCAGCGCAACGCGG	Rev SDM-*menT₄* D186A	pTRB621 cloning

Table 2.6. Plasmids used in this study

Plasmid	Genotype	Notes	Oligos	Source
pBAD30	pACYC184 derivative/p15A replicon	Ap^R, P_{BAD} promoter	N/A	[337]
pET-$MenT_{1\text{-}His}$	C-terminus His_6-tagged *M. tuberculosis* H37Rv *menT*$_1$, pETDuet-1 derivative	Ap^R, T7 promoter, *lac* operator	N/A	[98]
pLysSRARE (pRARE)	pACYC184 derivative, p15A replicon	Cm^R, *E. coli* rare tRNA	N/A	Novagen
pRLD30	*S. agalactiae* His_6-tagged *abiEi*, pTRB30 derivative, ColE1 replicon	Km^R, T5 promoter, *lac* operator	N/A	[83]
pRW50	RK2 replicon	Tc^R, promoterless *lacZ*-fusion vector	N/A	[338]
pSAT1-LIC	pBAT4 derivative; pMB1 replicon	Ap^R, T7 promoter, *lac* operator	N/A	[98]
pTA100	pQE-80L derivative with Sp^R cassette, ColE1 replicon	Sp^R, T5 promoter, *lac* operator	N/A	[31]
pPF656	*M. tuberculosis* H37Rv *menA*$_3$, pTA100 vector	Sp^R, T5 promoter, *lac* operator	N/A	[98]
pPF657	*M. tuberculosis* H37Rv *menT*$_3$, pBAD30 derivative	Ap^R, P_{BAD} promoter	N/A	[98]
pPF658	*M. tuberculosis* H37Rv *menA*$_4$, pTA100 derivative	Sp^R, T5 promoter, *lac* operator	N/A	[98]
pPF659	*M. tuberculosis* H37Rv *menT*$_4$, pBAD30 derivative	Ap^R, P_{BAD} promoter	N/A	[98]
pPF660	*M. tuberculosis* H37Rv *menA*$_2$, pTA100 derivative	Sp^R, T5 promoter, *lac* operator	N/A	[98]
pPF680	*S. agalactiae abiE* operon (P_{abiE}, *abiEi* R35A, C-terminus His_6-tagged *abiEii*), pQE-80L derivative	Ap^R, T5 promoter, *lac* operator	N/A	Fineran Lab
pTRB480	*M. tuberculosis* H37Rv *menT*$_2$, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1030 TRB1031	This study
pTRB481	*S. agalactiae abiEi*, pTA100 derivative	Sp^R, T5 promoter, *lac* operator	TRB1052 TRB1053	This study
pTRB482	*S. agalactiae abiEii*, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1054 TRB1055	This study
pTRB483	500 bp region upstream of *M. tuberculosis* H37Rv *menA*$_3$, pRW50 derivative	Tc^R	TRB1040 TRB1041	This study
pTRB484	500 bp region upstream of *M. tuberculosis* H37Rv *menA*$_4$, pRW50 derivative	Tc^R	TRB1042 TRB1043	This study

Table 2.6 continued. Plasmids used in this study

Plasmid	Genotype	Notes	Oligos	Source
pTRB485	500 bp region upstream of *M. tuberculosis* H37Rv *menA*$_2$, pRW50 derivative	Tc^R	TRB1044 TRB1073	This study
pTRB486	99 bp region upstream of *S. agalactiae abiEi*, pRW50 derivative	Tc^R	TRB1047 TRB1072	This study
pTRB491	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menA*$_3$, pSAT1-LIC derivative	Ap^R, T7 promoter, *lac* operator	TRB1018 TRB1019	This study
pTRB492	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT*$_3$, pSAT1-LIC derivative	Ap^R, T7 promoter, *lac* operator	TRB1020 TRB1021	This study
pTRB493	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menA*$_4$, pSAT1-LIC derivative	Ap^R, T7 promoter, *lac* operator	TRB1022 TRB1023	This study
pTRB494	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT*$_4$, pSAT1-LIC derivative	Ap^R, T7 promoter, *lac* operator	TRB1024 TRB1025	This study
pTRB496	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT*$_2$, pSAT1-LIC derivative	Ap^R, T7 promoter, *lac* operator	TRB1028 TRB1029	This study
pTRB517	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT*$_3$, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1120 TRB1121 TRB1122 TRB1124	This study
pTRB544	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT*$_4$, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1121 TRB1175 TRB1176 TRB1177 TRB1321 TRB1322	This study
pTRB549	pTA100 derivative with LIC site cloned in	Sp^R, T5 promoter, *lac* operator	TRB1461 TRB1462	This study
pTRB550	pBAD30 derivative with His_6-SUMO tag and LIC site cloned in	Ap^R, P_{BAD} promoter	TRB1460 TRB1462	This study
pTRB551	pBAD30 derivative with LIC site cloned in	Ap^R, P_{BAD} promoter	TRB1461 TRB1462	This study

Table 2.6 continued. Plasmids used in this study

Plasmid	Genotype	Notes	Oligos	Source
pTRB559	*M. tuberculosis* H37Rv *menT₃* K61A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1519 TRB1520	This study
pTRB560	*M. tuberculosis* H37Rv *menT₃* Q185A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1521 TRB1522	This study
pTRB561	*M. tuberculosis* H37Rv *menT₃* Q188A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1523 TRB1524	This study
pTRB562	*M. tuberculosis* H37Rv *menT₃* K189A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1525 TRB1526	This study
pTRB571	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT₃* Q185A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1521 TRB1522	This study
pTRB572	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT₃* Q188A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1523 TRB1524	This study
pTRB576	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT₃* K61A, pTRB550 derivative	Ap^R, P_{BAD} promoter	TRB1020 TRB1021	This study
pTRB577	N-terminus His_6-SUMO-tagged *M. tuberculosis* H37Rv *menT₃* K189A, pTRB550 derivative	Ap^R, P_{BAD} promoter	TRB1020 TRB1021	This study
pTRB591	*M. tuberculosis* H37Rv *menT₃* D80A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1676 TRB1677	This study
pTRB592	*M. tuberculosis* H37Rv *menT₃* D211A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1674 TRB1675	This study
pTRB593	*M.* tuberculosis H37Rv N-terminus His_6-SUMO-tagged *menT₃* D80A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1676 TRB1677	This study
pTRB594	*M. tuberculosis* H37Rv N-terminus His_6-SUMO-tagged *menT₃* D211A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1674 TRB1675	This study
pTRB597	*M. tuberculosis* H37Rv *menA₁*, pTRB549 derivative	Sp^R, T5 promoter, *lac* operator	TRB1681 TRB1682	This study

Table 2.6 continued. Plasmids used in this study

Plasmid	Genotype	Notes	Oligos	Source
pTRB617	*M. tuberculosis* H37Rv N-terminus His_6-SUMO-tagged *$menA_1$*, pTRB550 derivative	Ap^R, P_{BAD} promoter	TRB1699 TRB1700	This study
pTRB618	*M. tuberculosis* H37Rv *$menT_4$* S67A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1726 TRB1727	This study
pTRB619	*M. tuberculosis* H37Rv *$menT_4$* D69A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1728 TRB1729	This study
pTRB620	*M. tuberculosis* H37Rv *$menT_4$* K171A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1730 TRB1731	This study
pTRB621	*M. tuberculosis* H37Rv *$menT_4$* D186A, pBAD30 derivative	Ap^R, P_{BAD} promoter	TRB1732 TRB1733	This study
pTRB629	*M. tuberculosis* H37Rv N-terminus His_6-SUMO-tagged *$menT_1$*, pTRB550 derivative	Ap^R, P_{BAD} promoter	TRB1701 TRB1702	This study
pTRB631	*M. tuberculosis* H37Rv *$menT_3$* S78A, pTRB551 derivative	Ap^R, P_{BAD} promoter	TRB1678 TRB1021	This study

2.3.1. DNA isolation and purification

Plasmid DNA was extracted and purified using the Monarch Plasmid Miniprep Kit (New England Biolabs). PCR products and digested plasmids were separated and analysed by agarose gel electrophoresis, with the corresponding gel mix and TAE running buffer described in Table 2.2. Gels were visualised and imaged using a Gel Doc XR+ System (Bio-Rad) in conjunction with the Image Lab software package (Bio-Rad). DNA bands were excised from agarose gels and purified with the Monarch DNA Gel Extraction Kit (New England Biolabs). Digested PCR products were purified using the Monarch PCR and DNA Cleanup Kit (New England Biolabs). All purified DNA was eluted with MQ and either stored at -20 °C or kept at 4 °C for immediate use.

2.3.2. DNA manipulation and cloning

2.3.2.1. Polymerase chain reaction (PCR)

PCRs were typically performed for cloning purposes and were carried out using Q5 High-Fidelity DNA Polymerase (New England Biolabs) performed in a Mastercycler nexus X2 PCR thermocycler (Eppendorf). Genes were amplified from either genomic DNA, or sequence-verified plasmid DNA. PCR extension times were determined by target gene length, and annealing temperatures were calculated based on primer characteristics using the online NEB Tm calculator (New England Biolabs). Example reactions are detailed in Tables 2.7 and 2.8, with component volumes scalable depending on the desired PCR output. Reactions could be optimised by the addition of GC enhancer for GC-rich templates, and DMSO to reduce annealing temperatures. Where possible, template DNA consisted of either plasmid or genomic DNA diluted to approximately 2 ng/µl.

Where purified plasmid or genomic DNA was unavailable, amplicons were derived from bacteria by colony PCR. In these instances, cells were picked using a sterile pipette tip and added directly to reactions in place of template DNA. Reactions proceeded as normal, except for an additional 1 min of denaturation at 98 °C at the beginning of the PCR programme. The resulting PCR product was diluted 1:10 v/v in MQ and used as the template for a second PCR, using the same primer pair and reaction conditions, but omitting the extended denaturation step. This second reaction served to generate a cleaner PCR product, free of bacterial contamination and background amplification.

Table 2.7. Q5 PCR reaction components

Reaction component	Volume (µl)
MQ †	28.9 / 13.9
5x reaction buffer	10
5x GC enhancer (optional) †	- / 10
2 mM dNTPs	5
DMSO (optional) (5% final concentration) †	- / 5
10 mM Fwd primer	2.5
10 mM Rev primer	2.5
Template DNA (~2 ng/µl)	1
Q5 polymerase	0.1
	50

† MQ volume adjustable to accommodate optional use of DMSO or GC enhancer

Table 2.8. Q5 PCR reaction steps

Reaction step	Temperature (°C)	Time
Initial denaturation	98	30 sec
30 – 35 cycles		
· Denaturation	98	10 sec
· Annealing	50 - 72	30 sec
· Extension	72	30 sec/kb
Final extension	72	2 min
Hold	10	∞

2.3.2.2. Restriction digestion

Restriction digestion of DNA was performed using restriction enzymes (New England Biolabs, Thermo Fisher Scientific) singularly or in pairs, as per the manufacturers' instructions. In the case of double digestions, a compatible buffer for both enzymes was used. Alternatively, if a compatible buffer was unavailable, sample DNA was first digested with a single enzyme, gel-purified, digested again with the second enzyme, and gel-purified a final time. Typical reactions are described in Table 2.9 below; if multiple enzymes were used, 0.5 µl of each enzyme per 1 µg of sample DNA was added, with the MQ volume reduced accordingly. If using FastDigest (Thermo Fisher Scientific) or High-Fidelity (New England Biolabs) restriction enzymes, reactions were incubated for 30 min.

Table 2.9. Restriction digestion reaction

Reaction component	Volume (µl)	Reaction condition
MQ	26.5 - x	
DNA sample (1 µg)	x	1 - 2 h
10x enzyme buffer	3	37 °C
Enzyme	0.5	
	30	

2.3.2.3. DNA ligation

Ligation of restriction enzyme-digested vector and insert DNA was carried out using T4 DNA ligase (New England Biolabs) as per the manufacturer's instructions. Typical reactions are described in Table 2.10; molar insert:vector DNA ratios of 3:1 were predominantly used, as calculated by the online NEBioCalculator (New England Biolabs), with a standard vector concentration of 50 ng per reaction.

Table 2.10. DNA ligation reaction

Reaction component	Volume (µl)	Reaction condition
MQ	8.5 - x	
DNA vector	x	overnight, 16 °C ‡
DNA insert	x	10 min, 65 °C
10x T4 ligase buffer	1	
T4 DNA ligase	0.5	
	10	

‡ Alternatively, incubate 20 min at room temperature

2.3.2.4. Ligation independent cloning

Ligation independent cloning (LIC) is a quick and efficient means to anneal PCR products with a LIC-compatible vector independently of T4 DNA ligase[339]. In the context of this study, the pSAT1-LIC plasmid was routinely used for LIC cloning to generate constructs for recombinant protein expression. The pSAT1-LIC plasmid is a pBAT4 derivative which encodes a His_6-SUMO tag directly upstream of its LIC site[98,340]. This specific LIC site features a StuI cut site, which must first be digested to linearise the vector (Figure 2.1A). The 3' - 5' exonuclease activity of T4 DNA polymerase was then used to cut back linearised vector DNA. This activity can be controlled by the addition of specific NTPs; the addition of dTTP causes T4 DNA polymerase exonuclease activity to stall upon reaching a thymine, allowing the specific tailoring of 5' DNA overhangs and creation of the "LICed" vector (Figure 2.1A).

The LIC primers were designed to feature 5' tags homologous to the pSAT1-LIC LIC site, followed by a thymine, and finally the target gene sequence (Figure 2.1B). The translational start codon was omitted as the LIC process fuses the target gene to the pSAT1-LIC-encoded His_6-SUMO sequence, which encodes its own ribosome binding site (RBS) and start codon. DNA was amplified by Q5 PCR (Materials and Methods, *2.3.2.1*). The subsequent addition of T4 DNA polymerase and dATP cuts back the 3' amplified PCR product approximately 15 to 20 bases, with activity stalling at the introduced adenines (Figure 2.1B). This produces 5' ssDNA overhangs which are complimentary to those of the LIC-prepared vector (Figure 2.1B). After both the destination vector and PCR product were prepared by LIC, 2 µl of each LIC reaction were mixed by gentle pipetting and made up to a 10 µl total volume with MQ, then incubated at room temperature for 20 min. Due to the length of the complementary overhangs, when

mixed the PCR and vector DNA efficiently anneals without the need for T4 DNA ligase. This sample is then used to transform *E. coli* (Materials and Methods, *2.3.3*), which repairs and replicates the newly formed chimeric DNA plasmid. Typical LIC reactions for vector and insert DNA are described in Table 2.11.

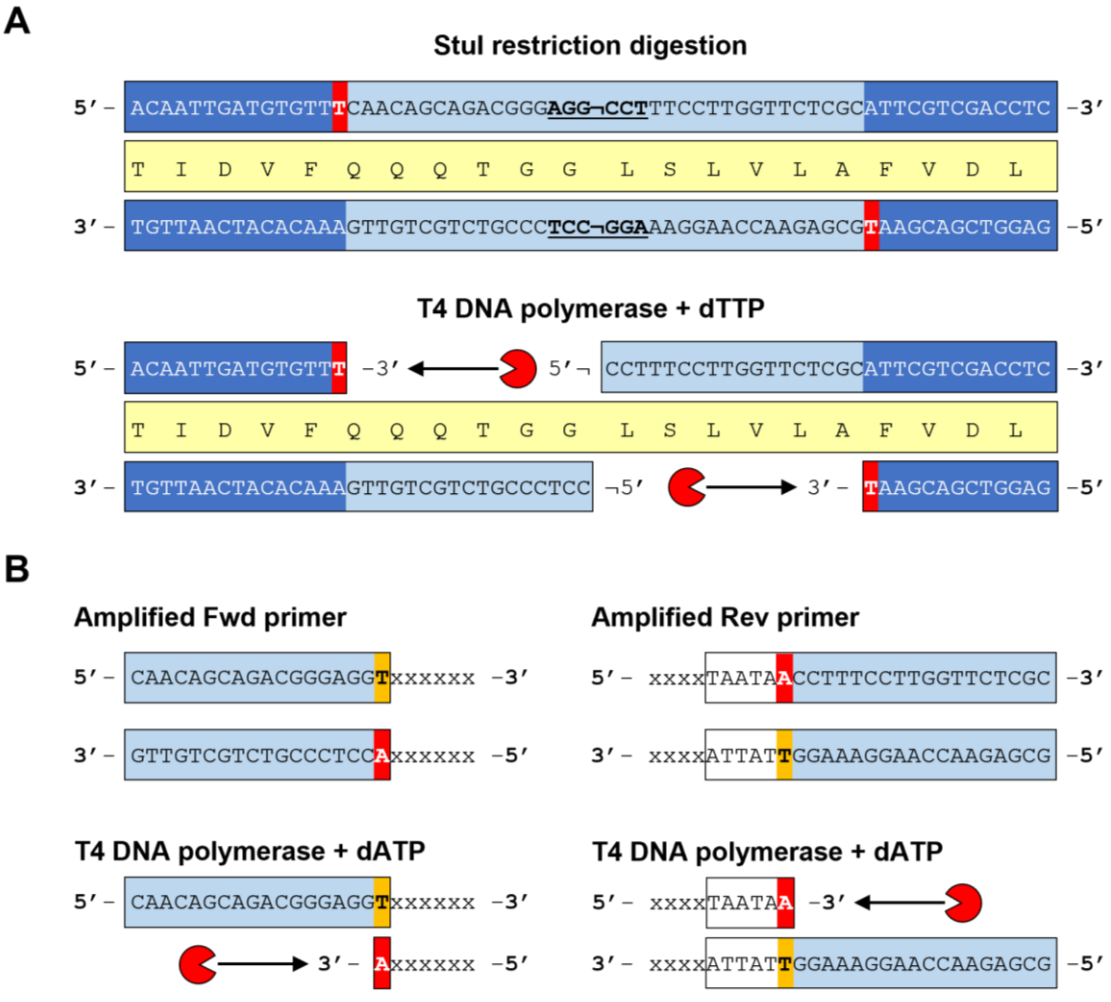


Figure 2.1. Schematic of LIC cloning. (A) The LIC vector contains a StuI restriction digestion site (underlined and in bold). The LIC site is coloured light blue; the AA sequence is coloured pale yellow and is situated between the complementary DNA sequences. Addition of T4 DNA polymerase and dTTP cuts back StuI-digested vector DNA until a thymine is reached (in bold, highlighted red). **(B)** LIC Fwd primers were designed with a 5′ tag homologous to the vector LIC site (light blue), plus an introduced thymine (in bold, coloured orange) immediately downstream of the tag and upstream of the ORF sequence (x). The LIC Rev primers were similarly designed with a 5′ tag matching the vector LIC site (light blue), plus an introduced double stop codon (first thymine in bold, coloured orange) immediately downstream of the tag and upstream of the ORF sequence (x). Primers were used to amplify DNA via PCR. Subsequent addition of T4 DNA polymerase and dATP cuts back DNA until an adenine (coloured red) is reached, generating 5′ overhangs complementary to linearised and T4 DNA polymerase-treated vector.

Table 2.11. Ligation independent cloning (LIC)

Reaction component	Volume (µl)	Reaction condition
LIC vector reaction		
MQ	12	
Gel-purified StuI-digested vector	25	
25 mM dTTP	5	30 min, 22 °C
10x 2.1 buffer (New England Biolabs)	5	20 min, 75 °C
100 mM DTT	2.5	
T4 DNA polymerase (New England Biolabs)	0.5	
	50	
LIC PCR reaction		
MQ	9.6	
PCR-purified PCR product	5	
25 mM dATP	2	30 min, 22 °C
10x 2.1 buffer (New England Biolabs)	2	20 min, 75 °C
100 mM DTT	1	
T4 DNA polymerase (New England Biolabs)	0.4	
	20	

The pTRB549, pTRB550, and pTRB551 LIC vectors were also used for LIC cloning. The pTRB550 plasmid was generated to function as a pBAD-based variant of pSAT1-LIC; it encodes the same His_6-SUMO-LIC site, amplified by PCR from pSAT1-LIC and cloned as an EcoRI/HindIII fragment into EcoRI/HindIII-digested pBAD30. The pTRB549 and pTRB551 vectors were created by amplifying the LIC site from pSAT1-LIC by PCR and cloning it as an EcoRI/HindIII fragment into EcoRI/HindIII-digested pTA100 and pBAD30, respectively. These two plasmids were used for LIC cloning of untagged genes and for expressing native, untagged proteins in functional assays.

2.3.2.5. Site-directed mutagenesis (SDM)

DNA was mutated using an in-house version of the Q5 Site-Directed Mutagenesis Kit (New England Biolabs). Non-overlapping inverse primers were designed using the online NEBaseChanger programme (New England Biolabs). Unless otherwise stated, SDM primers were designed to substitute the target native codon with one encoding an alanine. A Q5 PCR was then performed as described previously (Materials and Methods, *2.3.2.1*), using the target plasmid as template DNA and reducing the number of PCR cycles to 25. Finally, the PCR product was added to the SDM master mix, as described in Table 2.12 below, and incubated for 1 h at 37 °C, then at 16 °C for 1 h. A 10 µl aliquot of this reaction was then used to transform *E. coli* cells via heat-shock (Materials and Methods, *2.3.3*).

Table 2.12. SDM master mix

Reaction component	Volume (µl)
MQ	10
T4 DNA ligase buffer	2
Polynucleotide kinase (PNK) buffer	2
PCR product	2
ATP 25 mM	1
T4 DNA ligase	1
T4 PNK	1
DpnI	1
	20

2.3.2.6. DNA sequencing validation

All constructs were first verified by DNA sequencing prior to use, using the in-house DBS Genomics facility, Department of Biosciences, Durham University. Sample DNA constructs and custom sequencing primers were provided, and samples were analysed using a 3730 DNA Analyzer (Applied Biosystems). Sequencing data were analysed using Chromas 2.6 (Technelysium Pty Ltd), the NCBI Basic Local Alignment Search Tool (BLAST)[341], and the Translate tool via the ExPASy Bioinformatics Resource Portal.

2.3.3. *E. coli* heat-shock transformation

Competent *E. coli* cells were transformed by heat-shock treatment. Cells were first made competent via the chemically competent and ultra-competent methodology described below, then snap-frozen in liquid N_2 as 100 µl aliquots for storage at -80 °C. For heat-shock treatment, aliquots were first thawed on ice for 10 min, then quickly mixed by flicking the tubes. DNA was added, not exceeding 10% v/v of the competent cells. For ligation and LIC samples, the entire reaction was added to competent cells. For plasmid DNA, the amount of DNA used to transform cells was scaled according to copy number; for high copy number plasmids, 25 ng of DNA was added to competent cells, for low copy number plasmids, 100 ng was used. Cells and DNA were mixed by flicking, then incubated on ice for 15 min. Samples were heat-shocked for 45 sec at 42 °C in a pre-heated water bath, then left to rest on ice for 5 min. Cells were supplemented with 1 ml LB and allowed to recover for 1 h at 37 °C with shaking at 180 rpm. When transforming cells with pBAD30-based constructs, D-glu was added to the recovery media to repress gene expression. After 1 h, samples were centrifuged at 16,000 x *g* and the supernatant was discarded, leaving roughly 100 µl with which to resuspend the cell pellet. The resuspended cells were then spread on LBA plates containing the necessary antibiotics for plasmid selection, and D-glu where appropriate.

2.3.3.1. Chemically competent *E. coli* cells

E. coli strains were re-seeded at a 1:100 v/v ratio from overnight cultures into 250 ml flasks containing 25 ml of LB supplemented with 375 µl $MgCl_2$. These samples were grown at 37 °C with shaking at 180 rpm until an OD_{600} 0.4 – 0.6 was reached. Cells were transferred to a pre-chilled 50 ml centrifuge tube, incubated on ice for 1 h, then centrifuged at 4,200 x *g* for 10 min at 4 °C. The supernatant was removed and the cell pellet was resuspended in 10 ml ice-cold Solution A, then incubated on ice for 20 min. The sample was again centrifuged at 4,200 x *g* for 10 min at 4 °C, the supernatant was discarded, and the cell pellet was resuspended in 2 ml ice-cold Solution B. This resuspension was divided between pre-chilled 1.5 ml microcentrifuge tubes as 100 µl aliquots, then snap-frozen in liquid N_2 for storage at -80 °C.

2.3.3.2. Ultra-competent *E. coli* cells

Chemically ultra-competent *E. coli* DH5α cells were produced as previously described[342,343]. Briefly, on the morning that the protocol commenced, a single colony was used to inoculate a 250 ml flask containing 25 ml SOB medium, which was then grown at 37 °C with shaking at 180 rpm. Around 6 – 8 h later, bacteria were subcultured into three 1 L flasks, each containing 250 ml SOB. Each flask received a different volume of inoculum – 10 ml, 4 ml, or 2 ml – then all three were grown overnight at 18 °C with shaking at 180 rpm. The following morning, the OD_{600} of each flask was measured. Once one culture reached an OD_{600} 0.55, it was transferred to five pre-chilled 50 ml centrifuge tubes and stored on ice for 10 min, whilst the remaining two flasks were discarded. The chilled samples were centrifuged at 2,500 x *g* for 10 min at 4 °C and the supernatant was poured off. Samples were then re-centrifuged so that residual supernatant could be carefully removed by pipetting. Each cell pellet was resuspended in 16 ml ice-cold Inoue buffer by gentle pipetting in a 4 °C cold room, then centrifuged at 2,500 x *g* for 10 min at 4 °C to harvest cells. The supernatant was poured off and samples were centrifuged again to ensure the removal of all supernatant by pipetting. The resulting cell pellets were resuspended in 4 ml ice-cold Inoue buffer, combined into one tube, and 1.5 ml of DMSO was added. The bacterial suspension was gently mixed by swirling, then dispensed as 100 µl aliquots into pre-chilled 1.5 ml micro-centrifuge tubes and immediately snap-frozen in liquid N_2 for storage at -80 °C.

2.4. *E. coli* toxicity and antitoxicity assays

2.4.1. Endpoint viable count toxicity assays

In vivo toxicity assays in *E. coli* were performed as follows. *E. coli* DH5α were transformed with either the pBAD30 empty vector or the toxin constructs pTRB480 (*menT*$_2$), pPF657 (*menT*$_3$), pPF659 (*menT*$_4$) or pTRB482 (*abiEii*). Transformants were re-seeded at a 1:100 v/v ratio from overnight cultures into flasks containing LB supplemented with Ap and D-glu, and grown at 37 °C with shaking at 180 rpm to mid-log phase. Samples were normalised to an OD_{600} 1.0, serially diluted in PBS, then spotted on LBA or M9A plates supplemented with Ap and either D-glu or L-ara. Plates were incubated at 37 °C overnight (LBA) or for 48 h (M9A), then imaged and colonies counted to determine colony forming units (CFU)/ml.

Antitoxins were similarly tested in *E. coli* to assess for any growth defect caused by their expression. *E. coli* DH5α were transformed with either the pTA100 empty vector or the

antitoxin constructs pPF660 (*menA*$_2$), pPF656 (*menA*$_3$), pPF658 (*menA*$_4$) or pTRB481 (*abiEi*), then grown and tested as above. Antitoxin strains were supplemented with Sp, with or without IPTG. Data presented are the mean and standard deviation from three independent experiments.

2.4.2. Endpoint viable count antitoxicity assays

In vivo antitoxicity assays testing cognate or non-cognate antitoxins in *E. coli* were performed as follows. *E. coli* DH5α were co-transformed with pBAD30 empty vector, pTRB480, pPF657, pPF659 or pTRB482 (toxins), and either pTA100 empty vector, pPF660, pPF656, pPF658 or pTRB481 (antitoxins). Overnight cultures of transformants were re-seeded at a 1:100 v/v ratio into flasks containing fresh LB supplemented with Ap, Sp and D-glu, then grown at 37 °C to mid-log phase with shaking at 180 rpm. These samples were normalised to OD_{600} 1.0, then serially diluted in PBS and spotted onto LBA or M9A plates supplemented with Ap and Sp, and with or without IPTG, D-glu or L-ara. After an overnight (LBA) or 48 h (M9A) incubation at 37 °C, plates were imaged, colonies counted, and CFU/ml calculated. Data presented are the mean and standard deviation from at least three biological replicates.

2.4.3. Growth curves

E. coli strains carrying cognate toxins and antitoxins were assayed for growth inhibition and recovery in liquid media over a 24 h period. *E. coli* DH5α were first co-transformed with pBAD30 empty vector, pTRB480, pPF657, pPF659 or pTRB482 (toxins), and either pTA100 empty vector, pPF660, pPF656, pPF658 or pTRB481 (antitoxins). Single colony transformants were used to inoculate LB supplemented with Ap, Sp and D-Glu, and grown overnight. The following morning, overnight cultures were centrifuged at 16,000 x *g* and the supernatant was discarded, then the pellets were re-suspended in either LB (AbiEi-AbiEii) or M9M (MenA-MenT systems). Samples were subcultured into individual wells of a round-bottom 96-well microtest plate containing LB or M9M supplemented with Ap and Sp, to a starting OD_{600} 0.05 and a total volume of 200 µl. The microtest plate was then placed in a SPECTROstar Nano absorbance plate reader (BMG LABTECH) set to 37 °C with 200 rpm shaking. Samples were grown until mid-log phase, whereupon a multi-channel pipette was used to quickly add either D-glu to repress expression, IPTG and/or L-ara to induce expression, or sterile media to keep control sample volumes consistent. Cells were grown for an additional 24 h with OD_{600}

readings every 15 min. Data presented from kinetic growth curves are the mean and standard deviation from at least three biological replicates.

2.5. Promoter activity assays

Promoter activity assays were performed as described previously[344,345]. To assess promoter activity, *E. coli* DH5α were transformed with either the pRW50 empty vector or the *lacZ* reporter constructs pTRB485 (P_{menA_2}), pTRB483 (P_{menA_3}), pTRB484 (P_{menA_4}) or pTRB486 (P_{abiEi}). Transformants were re-seeded at a 1:100 v/v ratio from overnight cultures into LB supplemented with Tc, then grown at 37 °C with shaking at 180 rpm. When mid-log phase was reached, 80 µl of each sample was mixed with 120 µl of master mix (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 36 mM β-mercaptoethanol, 166 µl/ml lysozyme (from chicken egg white), 1.1 mg/ml ONPG, and 6.7% PopCulture Reagent (Merck Millipore)) in individual wells of a round-bottom 96-well microtest plate. This was then placed in a SPECTROstar Nano absorbance plate reader (BMG LABTECH) set to 30 °C with shaking at 500 rpm, wherein OD_{600} and OD_{420} readings were taken every 90 sec for 1 h. Data analysis was performed in the MARS Data Analysis software package (BMG LABTECH). The kinetic OD_{420} readings were converted into the slope of OD_{420} over time (OD_{420}/min). These values were multiplied by 5000 and divided by the OD_{600} reading from the first time point to generate Miller Units (mU). Data represent the mean and standard deviation from at least three biological replicates.

To assess autoregulation of the most active promoters, *E. coli* DH5α were co-transformed with either the pRW50 empty vector, pTRB484 (P_{menA_4}) or pTRB486 (P_{abiEi}), and either the pTA100 empty vector, pPF658 (*menA*$_4$) or pTRB482 (*abiEi*). Transformants were re-seeded from overnight cultures into LB supplemented with Tc, Sp, and with or without IPTG to induce antitoxin expression. Samples were then grown, tested, and analysed as described above. Data represent the mean and standard deviation from at least three biological replicates.

2.6. Recombinant protein expression

All expressions were performed using the same basic methodology. *E. coli* ER2566 or BL21(λDE3)Δ*slyD* expression strains were first transformed with the relevant expression plasmids. Single colonies were used to inoculate LB for overnight growth at 37 °C with 200 rpm shaking. At all stages of solid and liquid media growth, the relevant antibiotic selection

was included. For strains containing pBAD-based expression constructs, D-glu supplementation was only used to repress the expression of toxins that inhibited growth in *E. coli* functional assays. In these cases, D-glu was only included for transformations, and in the overnight starter culture step. Where this occurred, the following morning, prior to inoculation of the main expression media, overnight cultures were first centrifuged at 4200 x *g* for 15 min. The supernatant was discarded and the cell pellets were resuspended in fresh 2x YT. This wash step was repeated to remove any residual trace of D-glu, then the bacterial resuspension was re-seeded into 2x YT for expression.

For large scale expressions, overnight starter cultures were re-seeded at a 1:100 v/v ratio into 2 L baffled flasks containing 1 L 2x YT and antibiotics for selection. Cells were initially grown at 37 °C with 175 rpm shaking until an OD_{600} of 0.3 was achieved, then the temperature was reduced to approximately 25 °C until the expression culture reached an OD_{600} of 0.55. At this point, the temperature was turned down to 16 °C, the relevant inducing agents were added to express the target proteins, and the culture was grown overnight with shaking at 175 rpm.

2.6.1. $MenT_3$ selenomethionine incorporation

For $MenT_3$ selenomethionine (SeMet) incorporation, a starter culture of *E. coli* ER2566 pRARE pPF656 pTRB517 was first grown overnight in LB containing the relevant antibiotic selection, at 37 °C with 200 rpm shaking. D-glu supplementation was unnecessary due to the antitoxicity conferred by the $MenA_3$ antitoxin, and therefore omitted from all steps. The following morning, cells were centrifuged at 4200 x *g* for 15 min, washed and resuspended in M9M, then subcultured into 500 ml M9M plus antibiotic selection in 2 L baffled flasks to a starting OD_{600} of 0.075. Cells were grown at 37 °C with 175 rpm shaking to an OD_{600} of 0.6, whereupon cells were centrifuged at 4200 x *g* for 20 min, the supernatant was discarded, and pellets were resuspended in M9M. This sample was divided between separate 2 L baffled flasks containing fresh 500 ml M9M plus antibiotic selection, then shaken at 175 rpm for a further 1 h at 37 °C. Once an OD_{600} of 0.7 was reached, 12 ml of nutrient mix (Table 2.2) was added to each flask to promote feedback inhibition of methionine synthesis, followed by 250x SelenoMethionine Solution (Molecular Dimensions) to a final concentration of 40 µg/ml, and cells were left to incubate for 1 h at 20 °C. Finally, toxin and antitoxin expression were induced by the addition of L-ara and IPTG, respectively, and samples were left to grow overnight at 16 °C with 175 rpm shaking.

2.7. Protein purification

Unless otherwise stated below, all proteins were purified following the same procedure. Details for the buffers used are included in Table 2.2. Bacteria were harvested by centrifugation at 4200 x *g* and the pellets were resuspended in A500 buffer. Cells were lysed by sonication at 40 kpsi, then centrifuged (45,000 x *g*, 4 °C). The clarified lysate containing soluble protein was next passed over a HisTrap HP affinity chromatography column (Cytiva), washed for ten column volumes with A500, followed by ten column volumes of A100 buffer, then eluted directly onto a HiTrap Q HP anion exchange chromatography (AEC) column (Cytiva) with B100 buffer. The Q HP column was transferred to an Äkta Pure (Cytiva), washed with 3 column volumes of A100, then proteins were eluted using a gradient from 100% A100 to 100% C1000 buffer. Äkta fractions containing the target protein, as indicated by the AEC chromatogram peak, were first analysed and verified by SDS-PAGE. These were then pooled and incubated overnight at 4 °C with human sentrin/SUMO-specific protease 2 (hSENP2) to cleave the His_6-SUMO tag from the target protein (Figure 2.2). The next day, the sample was passed through a second HisTrap HP column and the flow-through containing untagged target protein was collected. This sample was concentrated and run over a HiPrep 16/60 Sephacryl S-200 size exclusion chromatography (SEC) column (Cytiva) connected to the Äkta Pure in Sizing buffer. Äkta fractions corresponding to the SEC chromatogram peak were analysed by SDS-PAGE, confirmed to contain target protein, then pooled and concentrated. Purified protein was either flash-frozen in liquid N_2 for storage at -80 °C or dialysed overnight at 4 °C into Crystal buffer (Table 2.2) for crystallographic studies. Crystallisation samples were quantified using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) and stored on ice, then either used immediately or flash-frozen in liquid N_2 for storage at -80 °C. Frozen crystallisation samples still formed usable crystals at least 15 months after storage.

$MenA_1$, $MenA_3$, and $MenA_4$ antitoxin proteins were purified with slight alterations to the standard procedure described above. For $MenA_1$, the SEC step was skipped; a previous expression which had yielded large amounts of protein (as indicated by AEC and SDS-PAGE) resulted in an almost total loss of protein during SEC. This was presumed to be due to the unstable nature of native unbound $MenA_1$ protein, and it being ill-suited to the HiPrep 16/60 Sephacryl S-200 SEC column used. Nevertheless, subsequent $MenA_1$ purifications that culminated with the second HisTrap HP column post-hSENP2 treatment were more than sufficient to purify and isolate $MenA_1$ protein.

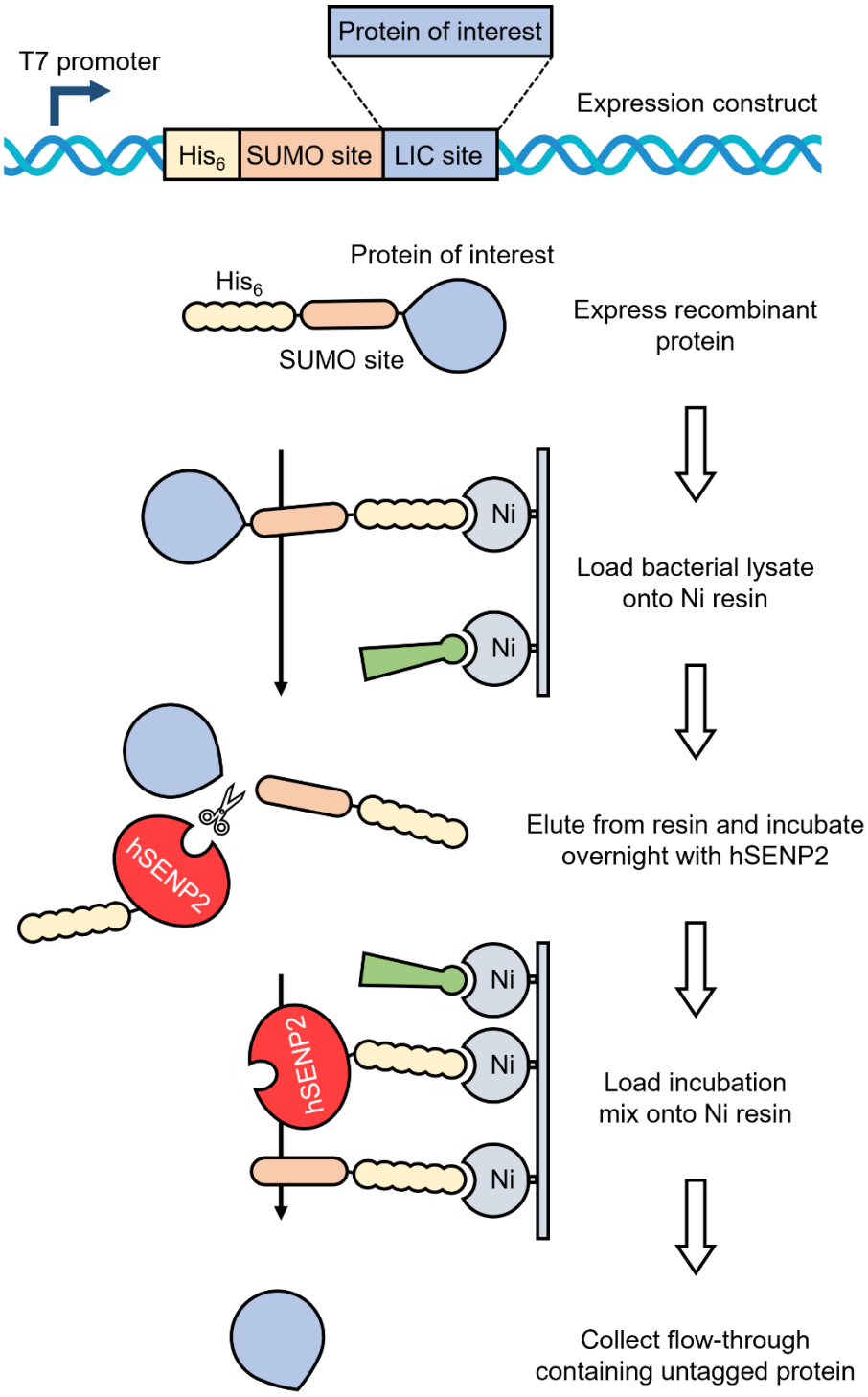


Figure 2.2. Schematic outlining the purification of His_6-SUMO fusion proteins using two-step affinity chromatography. To generate expression constructs, the protein of interest was fused via LIC to a cleavable His_6-SUMO domain. Following expression, the cell lysate was passed through a HisTrap HP Ni affinity chromatography column. The column was washed and all bound proteins, including background Ni-binding proteins, were eluted. The resulting sample was incubated overnight at 4 °C with His_6-tagged hSENP2 SUMO protease, which cleaves at the SUMO site to remove the His_6-SUMO tag from the target protein. The overnight incubation mix was then passed down a second HisTrap HP column; whilst the contaminant proteins and hSENP2 retain Ni affinity, the newly untagged target protein is isolated and collected in the flowthrough.

$MenA_3$ proved considerably more difficult to purify. During the first purification attempt, as soon as clarified lysate was passed down the first HisTrap HP column, $MenA_3$ protein began precipitating out of solution. This phenomenon continued to occur during all stages of purification, regardless of the buffer used. To rescue residual $MenA_3$ protein for subsequent purification steps, samples were centrifuged at 13,000 x *g* for 2 min and the supernatant containing remaining soluble $MenA_3$ was removed. To remedy $MenA_3$ protein precipitation issues, two measures were employed. Firstly, buffer optimisation was carried out, which identified that $MenA_3$ appeared partially preferent for high salt buffers, either NaCl of KCl, of approximately 0.8 M. Therefore, further purifications utilised high salt (0.8 M) buffers otherwise equivalent to those used in the standard purification procedure. Secondly, the number of purification steps were reduced. The resulting procedure employed a HisTrap HP column elution gradient performed on the Äkta Pure, using a gradient from 100% A500 (0.8 M NaCl) to 50% B100 (0.8 M NaCl). Fractions containing the protein peak were dialysed overnight at 4 °C into A500 (0.8 M NaCl) with hSENP2. The following morning, the sample was passed over a second HisTrap HP column and the flow-through fractions containing untagged $MenA_3$ protein were collected and verified by SDS-PAGE.

$MenA_4$ purification was relatively simpler to adapt; after loading hSENP2-cleaved $MenA_4$ onto the second HisTrap HP column, $MenA_4$ was absent from the flow-through, indicating retained Ni affinity despite removal of the His_6SUMO tag. As a result, a HisTrap HP column elution gradient was performed on the Äkta Pure, using a gradient from 100% A500 to 50% B100 (0.5 M NaCl). This was then carried forward into SEC as per the standard procedure described above.

2.8. Protein crystallisation

Native and SeMet-derivatised $MenT_3$ SEP and native $MenT_3$ were concentrated to 12 mg/ml, and native $MenT_4$ to 6 mg/ml, in Crystal buffer. Initial crystallisation screens were performed using a Mosquito Xtal3 robot (SPT Labtech), set to either 200:100 nl and 100:100 nl ($MenT_3$ SEP and $MenT_3$) or 100:50 nl and 100:100 nl ($MenT_4$) protein:condition sitting drops. After initial screening, $MenT_3$ SEP and $MenT_3$ protein samples formed thick, six-sided needles in condition G5 (0.2 M calcium acetate hydrate, 0.1 M Tris pH 8.5 and 25% w/v PEG 2000 MME) of Clear Strategy II HT-96 (Molecular Dimensions). $MenT_4$ formed thin, six-sided needles in the same condition as $MenT_3$. This condition was first optimised to enhance protein crystal formation. To harvest, 20 µl of condition reservoir was added to 20 µl of Cryo buffer (Table

2.2) and mixed quickly by vortexing; this mixture was then added to the protein crystal drop at a 1:1 v/v ratio. After the addition of cryoprotectant, crystals were immediately extracted using a nylon loop and transferred to a uni-puck stored in liquid N_2.

$MenA_1$:$MenT_1$ and $MenT_1$ protein samples were also concentrated to 12 mg/ml in Crystal buffer and crystallisation screens were performed as described for $MenT_3$ above. After initial screening, $MenA_1$:$MenT_1$ formed a single large cuboid crystal in condition F1 (0.2 M sodium fluoride, 0.1 M Bis-Tris propane pH 6.5 and 20% w/v PEG 3350) of PACT Premier Eco HT-96 (Molecular Dimensions), and $MenT_1$ formed thin square crystals in conditions E8 (1.8 M lithium sulphate and 0.1 M Tris pH 7.5) and G8 (1.8 M lithium sulphate and 0.1 M Tris pH 8.5) of Clear Strategy II Eco HT-96 (Molecular Dimensions). Crystals were harvested direct from the screen conditions, as described above.

2.9. X-ray data collection and structure determination

Diffraction data were collected at Diamond Light Source on beamlines I04 ($MenA_1$:$MenT_1$ and $MenT_3$ SEP), I03 ($MenT_3$ SEP SeMet-derivatised) and I24 ($MenT_1$, $MenT_3$ and $MenT_4$) (Tables 5.1 and 5.2). Two 360° datasets were collected for $MenA_1$:$MenT_1$ at wavelength 0.9795 Å and merged using iSpyB (Diamond Light Source). Single 720° datasets were collected for $MenT_3$ and $MenT_1$ at 0.9795 Å, and single 360° datasets were collected for $MenT_4$ at 0.9781 Å. For $MenT_3$ SEP, single 360° datasets were collected for native $MenT_3$ SEP at wavelength 0.9795 Å. Two 360° datasets from $MenT_3$ SeMet-derivatised crystals measured at the selenium peak (0.9793 Å) were merged using iSpyB, and additional SeMet-derivatised datasets were collected at selenium high remote (0.9641 Å) and inflection (0.9795 Å) wavelengths. Diffraction data were processed with XDS[346,347], and then AIMLESS from CCP4 was used to corroborate the spacegroups[348] (Tables 5.1 and 5.2).

The crystal structure of $MenT_3$ SEP was solved by multi-wavelength anomalous diffraction (MAD) by providing the SHELX suite in CCP4 with the native and three anomalous $MenT_3$ SEP datasets[349]. The $MenT_3$ structure was determined by molecular replacement (MR) via PHASER[350], using the $MenT_3$ SEP structure as a search model. The crystal structures of $MenA_1$:$MenT_1$ and $MenT_4$ were solved *ab initio* using ARCIMBOLDO[351]. The $MenT_1$ structure was also solved by PHASER MR[350], with the $MenA_1$:$MenT_1$ structure used as a search model. All solved crystal structures were further built using BUCCANEER and REFMAC in CCP4[348,352,353], then iteratively refined and built using PHENIX and COOT, respectively[354–357]. The quality of the final model was assessed using COOT and the wwPDB validation

server[357,358]. Structural figures, including alignments and superpositions, were generated using PyMol (Schrödinger). Residue conservation was calculated by ConSurf[359]; comparisons against models within the PDB were performed using DALI[360]. Structure modelling was performed using Phyre2[361], AlphaFold[362], and HADDOCK2.4[363].

2.10. Analytical SEC

A Superdex 200 Increase 5/150 GL SEC column (S200-I; Cytiva) connected to an Äkta Pure (Cytiva) was pre-equilibrated with two column volumes (6 ml) of Sizing buffer. Before and between experiments, the 10 µl sample loading loop was washed with 500 µl of filtered MQ, followed by 500 µl of Sizing buffer, using a 500 µl Hamilton syringe. Protein samples were prepared in 50 µl volumes containing 10 µM of protein in Sizing buffer. Mixed protein samples were prepared at a 1:1 molar ratio, containing 10 µM of each protein made up to 50 µl in Sizing buffer. Samples were incubated for 30 min at room temperature, then loaded into the equilibrated 10 µl loop using a 50 µl Hamilton syringe. Samples were injected directly onto the S200-I by running 1.3 column volumes (4 ml) of Sizing buffer through the sample loop at 0.180 ml/min; the 0.3 column volumes of extra buffer served to pre-equilibrate the S200-I for the next run. The UV absorbance at 280 nm was measured and presented as milli-arbitrary units (mAU), corresponding to the relative protein concentration. Data presented are representative of at least three biological replicates.

2.11. Mass spectrometry analysis

Purified protein samples were dialysed overnight at 4 °C into 10 mM ammonium bicarbonate using a Pur-A-Lyzer Mini Dialysis Kit (Sigma-Aldrich), at an approximately 1:1000 v/v sample:buffer ratio. Final samples were prepared in 130 µl volumes at a protein concentration of approximately 0.5 mg/ml. These were then sent for in-house analysis within the Department of Chemistry, Durham University, on an Acquity Ultra-High Performance Liquid Chromatograph (Waters) via positive electrospray ionisation time-of-flight mass spectrometry (ES+ TOF MS). Mass/charge ratios were obtained and data was analysed in-house using the RemoteAnalyzer software package (SpectralWorks).

2.12. Microscopy

E. coli DH5α were transformed with either the pBAD30 empty vector or the toxin constructs pTRB480 (*menT*$_2$), pPF657 (*menT*$_3$), pPF659 (*menT*$_4$) or pTRB482 (*abiEii*). Overnight cultures were centrifuged at 16,000 x *g* and the cell pellets were resuspended in fresh M9M. These samples were used to inoculate flasks of fresh M9M supplemented with Ap and L-ara to a starting OD_{600} 0.3. After 6 h growth shaking at 180 rpm in 37 °C, 250 µl of cells were removed and centrifuged at 8,000 x *g* for 2 min, then resuspended in 100 µl of 10 µg/ml Hoechst 33342 (Thermo Fisher Scientific) in 1x PBS. A 1% agarose gel pad was created using 1x PBS as solvent, which was then added to the centre of a concave glass microscope slide. Cells were added as 20 µl drops onto the surface of corresponding agarose pads, then sealed with coverslips. Samples were immediately taken for imaging on a Leica TCS SP5 II (Leica Microsystems) laser-scanning confocal microscope on a Leica HCX PL APO lambda blue 40X 1.52 OIL UV objective, including 2.5x digital zoom. Brightfield and fluorescence data at λ_{exc} = 405 nm, λ_{em} = 450 nm were captured. Images were analysed using the LAS X (Leica Microsystems) software package. The presented results are representative of data collected from two biological replicates.

2.13. Cell-free expression assay

Cell-free expression assays were performed as previously described[133]. Briefly, protein synthesis from dihydrofolate reductase (DHFR) template DNA (P0ABQ4, New England Biolabs) using *in vitro* transcription/translation coupled assays (PURExpress *In Vitro* Protein Synthesis Kit, New England Biolabs) was assessed in the absence and presence of increasing concentrations of purified MenT toxin or MenA antitoxin proteins. These assays were performed according to the manufacturer's instructions, with reactions incubated for 2 h at 37 °C. Protein synthesis was assessed by SDS-PAGE and densitometry analyses performed using the Image Lab software package (Bio-Rad). Data presented are the mean and standard deviation from three biological replicates.

2.14. Data analysis

The data analysis software used for specific experiments is cited in the relevant Materials and Methods sections. All other data analysis not otherwise mentioned, including statistical

analysis and data formatting, was performed in Microsoft Excel. Figures were generated using Microsoft PowerPoint and the online BioRender software.

3. Functional characterisation of three *M. tuberculosis* TA systems

3.1. *M. tuberculosis* encodes three putative type IV TA systems

At the project outset, research efforts focused on three putative type IV TA systems from *M. tuberculosis*: $MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$. These systems were first identified by Dy *et al*. and Sala *et al*. in 2014[11,83]. All three MenT toxins feature a conserved NTase-like domain, annotated as DUF1814, which is shared by the type IV AbiEii toxin from *S. agalactiae*[11,83] (Figure 3.1). DUF1814 proteins are found widespread among approximately 3000 bacterial, archaeal, and fungal genomes, though not all DUF1814 proteins are associated with putative antitoxins[83]. NTase-like DUF1814 proteins contain four conserved motifs: N-terminal motifs I and II are found in DNA polymerase β, where they are associated with metal ion coordination for nucleotide binding and transfer[83]; C-terminal motif III is similar to that found in tRNA NTases which add CCA to the 3' terminus of immature tRNAs, and is thought to mediate base-stacking of incoming nucleotides[364]; while C-terminal motif IV is exclusive to DUF1814 proteins, where it is proposed to form a catalytic site with motif III[83]. Notably, motifs I, II and IV are essential for AbiEii toxicity and NTase activity[83].

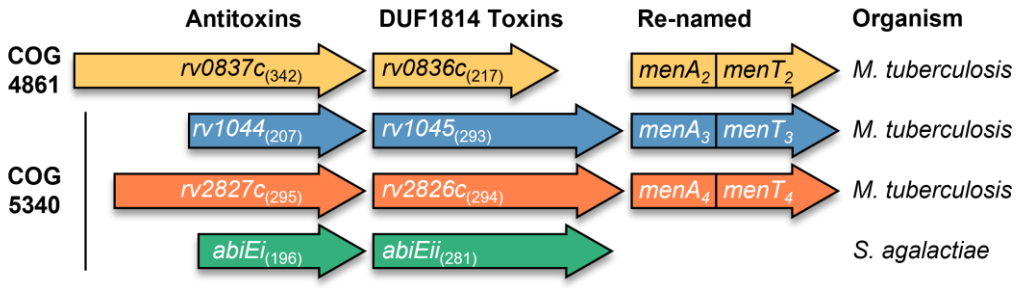


Figure 3.1. *M. tuberculosis* encodes at least three TA systems with DUF1814 NTase-like toxins. Graphical representation of the initial three putative *M. tuberculosis* TA systems containing DUF1814 NTase-like toxin genes, along with the homologous *S. agalactiae* AbiEi-AbiEii system, scaled according to protein AA composition (detailed in parentheses). The assigned COG group of each antitoxin is denoted to the left. Each TA system has been allocated a specific colour, retained throughout the manuscript. Revised nomenclature and the host organism are depicted to the right.

Though not always the case, DUF1814 proteins are often found located adjacent to COG5340 or COG4861 genes, with protein-association predictions suggesting a functional link between these protein families[83]. Correspondingly, the DUF1814 *menT* toxin genes are also encoded adjacent to and downstream of COG5340 and COG4861 putative antitoxin genes in *M. tuberculosis*. $MenA_3$ and $MenA_4$ belong to the COG5340 family of predicted transcriptional

regulators, as does the type IV AbiEi antitoxin, while $MenA_2$ is a COG4861 protein family member[11,83] (Figure 3.1). Although previously annotated as uncharacterised, COG4861 proteins have recently been updated to the CL0578/AbiEi superfamily alongside COG5340, and are now similarly classified as predicted transcriptional regulators. This putative role, alongside their genetic location upstream of DUF1814 genes, is consistent with a potential function for the COG5340 and COG4861 MenA proteins as autoregulatory antitoxins, and supports the hypothesis that the MenA-MenT gene pairs are TA systems. Indeed, the COG5340 AbiEi antitoxin has previously been shown to bind and repress transcription from the *abiEi-abiEii* promoter, similar to type II TA autoregulation[84].

The AbiEii, $MenT_3$, and $MenT_4$ DUF1814 toxins are comparable in terms of AA content, consisting of between 281 to 294 AAs in their annotated sequences. In contrast, $MenT_2$ is somewhat shorter, comprising only 217 AAs. Despite belonging to a conserved protein family, sequence identity between the DUF1814 toxins is relatively poor. Following pairwise protein sequence alignments with AbiEii (EMBOSS Stretcher), the highest MenT sequence identity is only 25.3%, shared with $MenT_4$. The MenA antitoxins exhibit similarly poor sequence identity to AbiEi, where the highest score is 23.6%, shared with $MenA_3$. The variability in antitoxin size compared to the cognate toxin is also interesting. AbiEi and $MenA_3$ are both approximately a third smaller than their cognate toxins, whilst $MenA_4$ is almost identical in size to $MenT_4$, and *menA₂* encodes a considerably larger antitoxin than its cognate $MenT_2$ toxin (Figure 3.1). The differences between these putative TA systems likely reflect evolutionary divergence in structure and function.

3.2. Production of toxin and antitoxin expression constructs

$MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$ were identified as putative type IV TA systems based on conserved functional domains shared with the *S. agalactiae* AbiEi-AbiEii system. However, the *M. tuberculosis* systems remained functionally uncharacterised. Therefore, it was decided to first test the phenotypic effects of toxin and antitoxin expression and investigate TA activity *in vivo*. The *E. coli* DH5α strain was chosen for phenotypic testing, with its safety profile, rapid growth, and relative ease to culture compared to *M. tuberculosis* presenting considerable technical benefits to the experimental aims. Potential issues regarding the use of *E. coli* related to the documented biological differences between *E. coli* and *M. tuberculosis*, ranging from cell cycle regulation to the mechanics of protein synthesis[365–367]. However, *E. coli* still provided considerable biological relevance; the

conservation of homologous TA systems among different bacterial organisms, including in *E. coli* and *M. tuberculosis*, is well established[11,12], and numerous prior studies have characterised the activity of *M. tuberculosis* TA systems in *E. coli*[368–372].

As such, the respective *menA*/*menT* antitoxin/toxin genes were cloned into separately inducible *E. coli* plasmids. Firstly, the *menA* and *menT* genes were amplified by PCR from *M. tuberculosis* H37Rv genomic DNA. Because the effects of toxin expression were unknown, and basal expression from leaky promoters might prove prematurely deleterious to cell growth, plasmid pBAD30 was chosen as a destination vector for toxin genes[337]. This low-copy number plasmid contains the P_{BAD} promoter, which allows for tightly controllable expression via L-ara induction and D-glu repression. Plasmids pPF657 and pPF659 were kindly provided by the Fineran lab, and featured the *menT*$_3$ and *menT*$_4$ genes, respectively, cloned as MfeI/XmaI fragments into EcoRI/XmaI-digested pBAD30. The *menT*$_2$ gene was cloned as an EcoRI/HindIII fragment into EcoRI/HindIII-digested pBAD30, to generate pTRB480. To express antitoxins, genes were cloned into pTA100, an IPTG-inducible pQE-80L derivative[31]. Plasmids pPF660, PF656, and pPF658 were also provided by the Fineran lab, and contained the *menA*$_2$, *menA*$_3$, and *menA*$_4$ antitoxin genes, respectively, cloned as MfeI/XmaI fragments into EcoRI/XmaI-digested pTA100.

In order to thoroughly assess toxicity, as well as provide a positive control of TA activity, the well-characterised and homologous AbiEi-AbiEii TA module from *S. agalactiae* was also tested. The *abiEi* gene was amplified from pRLD30, and *abiEii* from pPF680; both PCR products were separately cloned as EcoRI/HindIII fragments into EcoRI/HindIII-digested pTA100 and pBAD30, respectively, producing pTRB480 and pTRB482.

3.3. MenT toxins do not inhibit *E. coli* growth in nutrient-rich media

The first aim of this project was to test the putative MenA-MenT TA systems in *E. coli* for characteristic TA activity. Endpoint viable count toxicity and antitoxicity assays were performed as described in Materials and Methods, *2.4.1* and *2.4.2*. *E. coli* DH5α were transformed with the constructs described in *Section 3.2*, and used to assess the effects of ectopic antitoxin and toxin expression on bacterial growth in nutrient-rich LB media.

Typically, antitoxins do not inhibit growth, therefore $MenA_2$, $MenA_3$, $MenA_4$, and AbiEi were first tested to confirm that they impart no aberrative growth phenotypes (Figure 3.2). As expected, cell growth was unaffected by antitoxin expression alone, with viable counts and CFU/ml comparable to uninduced samples (Figure 3.2). Interestingly, induction of $MenA_4$

expression slightly altered colony morphology, producing marginally larger colonies with irregular, almost "blurred" outlines. Next, the effects of toxin expression on *E. coli* growth were tested (Figure 3.3). On LBA plates, no reductions in viable counts were observed as a result of MenT toxin expression. In contrast, induction of AbiEii caused a roughly 6-$\log_{10}$ drop in CFU/ml, corroborating previous findings and reflecting its potent toxicity in *E. coli*[83]. Unique among the MenT toxins, under these conditions $MenT_3$ expression appeared to attenuate growth, represented by the formation of visibly smaller and paler colonies (Figure 3.3). However, despite their weaker appearance, overall colony numbers were still comparable to uninduced $MenT_3$.

The effects of cognate toxin and antitoxin co-expression were next tested (Figure 3.4). Dual-plasmid *E. coli* DH5α strains were generated containing the cognate toxins and antitoxins on separately inducible plasmids, then assayed under a variety of induction conditions. As expected, and corroborating earlier data, expression of the three MenT toxins failed to affect growth in LB. This consequently rendered the results of MenA antitoxin co-expression mostly redundant. Interestingly, however, in the presence of the $MenA_3$ antitoxin, even with $MenA_3$ uninduced, the attenuated growth phenotype observed from lone $MenT_3$ expression was lost (Figures 3.3 and 3.4). This suggests that the $MenA_3$ construct is sufficiently leaky to neutralise mild $MenT_3$ toxicity. Unsurprisingly, AbiEii expression inhibited growth whilst AbiEi co-expression maintained growth, indicating that AbiEi is antitoxic and confirming that AbiEi-AbiEii functions as a TA system in *E. coli*[83] (Figure 3.4).

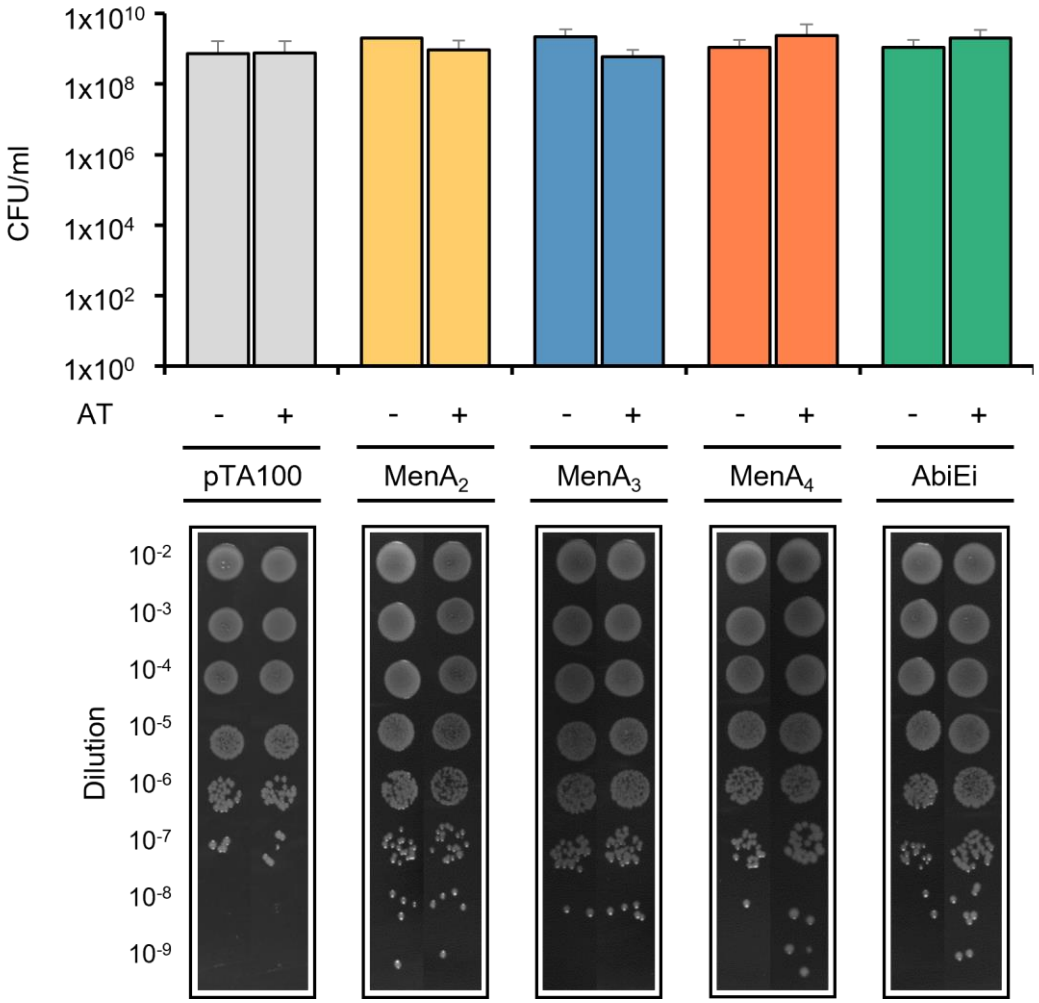


Figure 3.2. MenA antitoxin expression does not affect *E. coli* growth in LB. Endpoint viable count toxicity assays of *E. coli* DH5α transformed with the pTA100 vector, pPF660 (*menA*$_2$), pPF656 (*menA*$_3$), pPF658 (*menA*$_4$) or pTRB481 (*abiEi*). Overnight cultures were re-seeded into fresh LB supplemented with Sp and grown to mid-log phase. Samples were serially diluted and spotted on LBA plates containing Sp alone, or with IPTG to induce antitoxin expression. Plates were incubated at 37 °C overnight. The following morning, plates were imaged and colonies counted to determine CFU/ml. "AT" = antitoxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

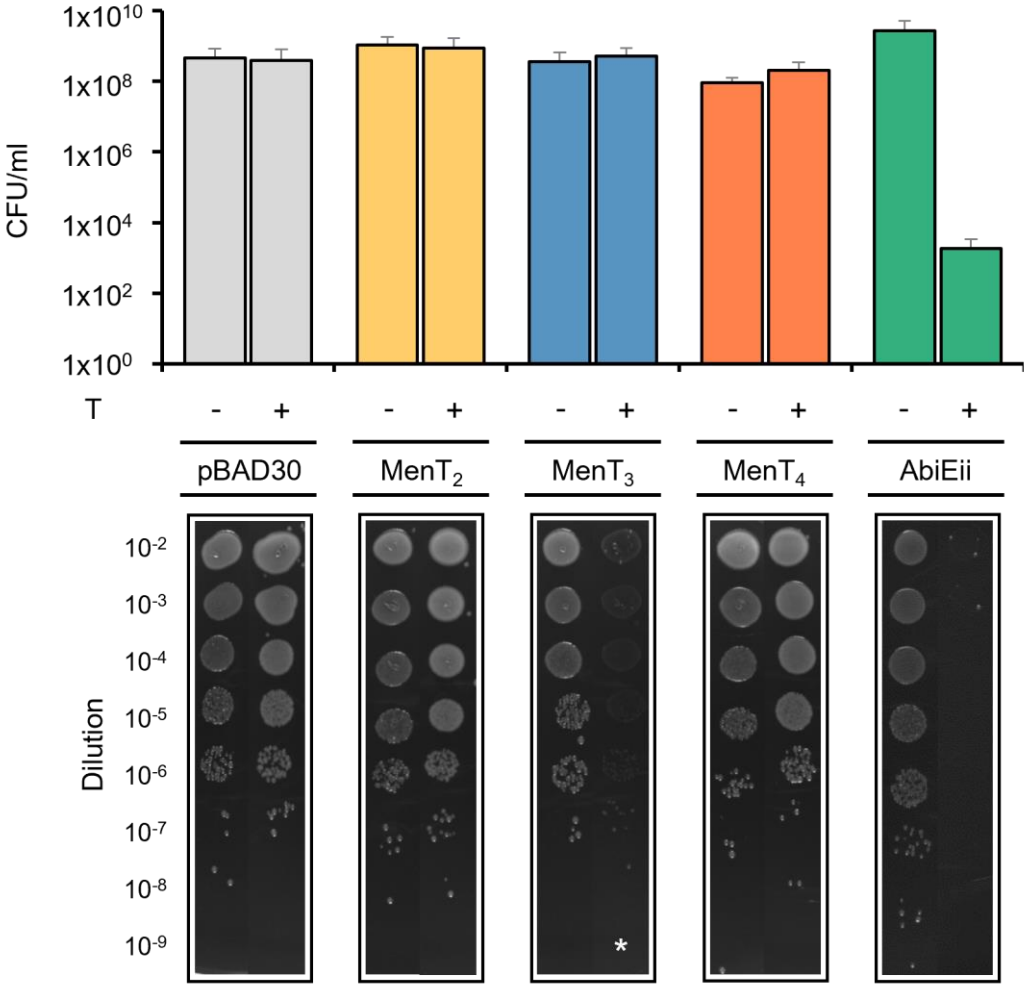


Figure 3.3. MenT toxins have varying effects on *E. coli* cell growth in nutrient-rich media. Endpoint viable count toxicity assays of *E. coli* DH5α transformed with the pBAD30 vector, pTRB480 (*menT*$_2$), pPF657 (*menT*$_3$), pPF659 (*menT*$_4$) or pTRB482 (*abiEii*). Overnight cultures were re-seeded into fresh LB supplemented with Ap and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on LBA plates containing Ap and either D-glu or L-ara, for repression or induction of gene expression, respectively. Plates were incubated at 37 °C overnight. The following morning, plates were imaged and colonies counted to determine CFU/ml. "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates). The white asterisk highlights a growth phenotype associated with MenT$_3$ expression which is only partially represented by the spot assay images: whilst growth is visibly attenuated compared to the uninduced MenT$_3$ sample, colonies are still formed, but are noticeably smaller and paler.

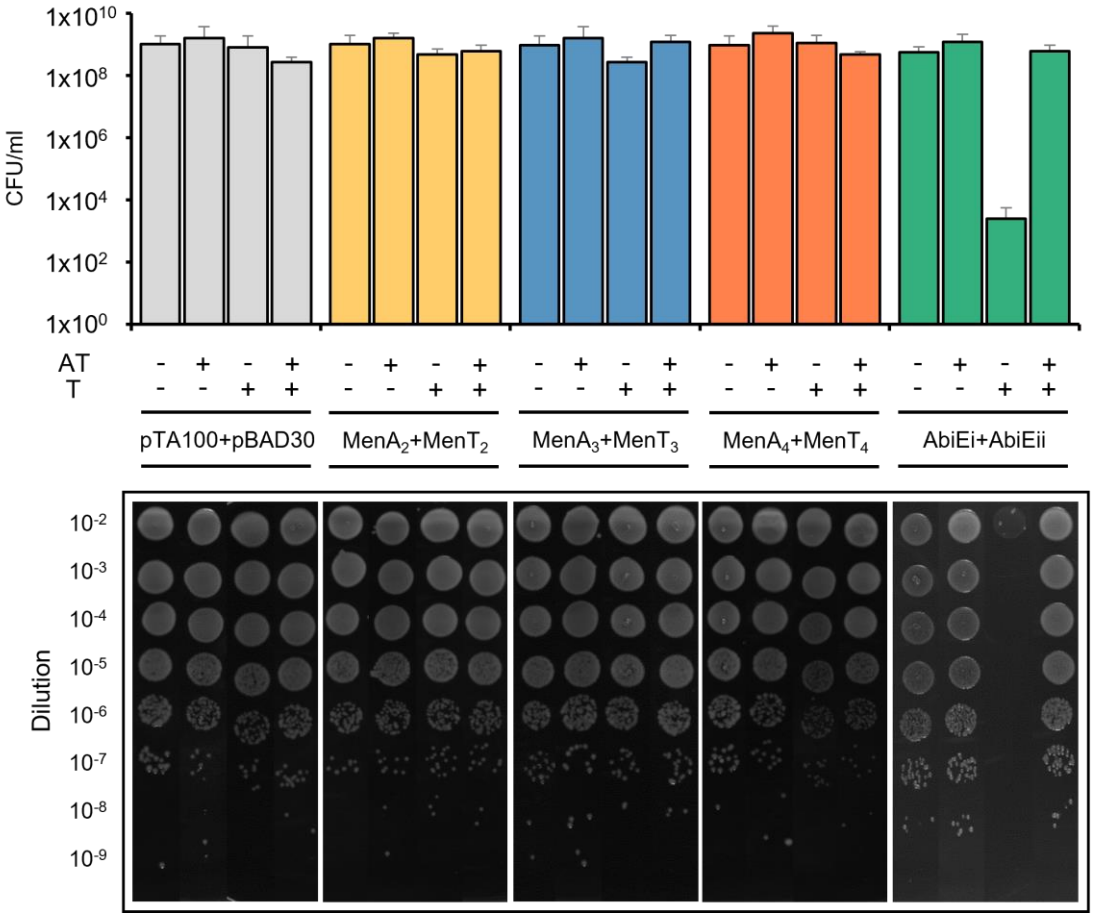


Figure 3.4. MenA-MenT systems do not exhibit TA activity in nutrient-rich conditions. Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with either empty vectors or the cognate toxins and antitoxins on separately inducible plasmids (pTA100 + pBAD30, pPF660 (*menA*$_2$) + pTRB480 (*menT*$_2$), pPF656 (*menA*$_3$) + pPF657 (*menT*$_3$), pPF658 (*menA*$_4$) + pPF659 (*menT*$_4$), or pTRB481 (*abiEi*) + pTRB482 (*abiEii*)). Overnight cultures were re-seeded into fresh LB supplemented with Ap, Sp and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on LBA plates containing Ap and Sp, and with or without D-glu, L-ara and IPTG for repression of toxin expression, induction of toxin expression, and induction of antitoxin expression, respectively. Plates were incubated at 37 °C overnight. The following morning, plates were imaged and colonies counted to determine CFU/ml. "AT" = antitoxin, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

3.4. MenA$_3$-MenT$_3$ and MenA$_4$-MenT$_4$ function as TA systems in *E. coli*

The absence of MenT-mediated growth inhibition in LB media was perplexing. Whilst the DUF1814 AbiEii toxin demonstrated marked toxicity, the only perceptible phenotype associated with MenT toxins was a minor impairment to colony formation conferred by MenT$_3$ expression. Several factors could account for these observations: the putative *M. tuberculosis* MenT toxins might not be functional in *E. coli,* they may not function as TA

systems at all, or an overlooked aspect of the experimental assays might be to blame. The activation of specific TA systems has often been linked with stress conditions such as nutrient starvation (See, *1.3.3*). It was therefore hypothesised that the lack of toxin activity might be caused by the nutrient-rich media in which they were tested. The use of minimal media in TA growth assays has been reported previously[81,373–375]; accordingly, toxicity and antitoxicity assays were repeated, but with M9A minimal media replacing LBA to reproduce more stressed, nutrient-sparse conditions for bacterial growth. Expression of the putative MenA antitoxins again did not inhibit growth (Figure 3.5), consistent with earlier results, although similar to the phenotype observed in LB media, $MenA_4$ expression led to a slightly enlarged and irregular colony morphology (Figure 3.5).

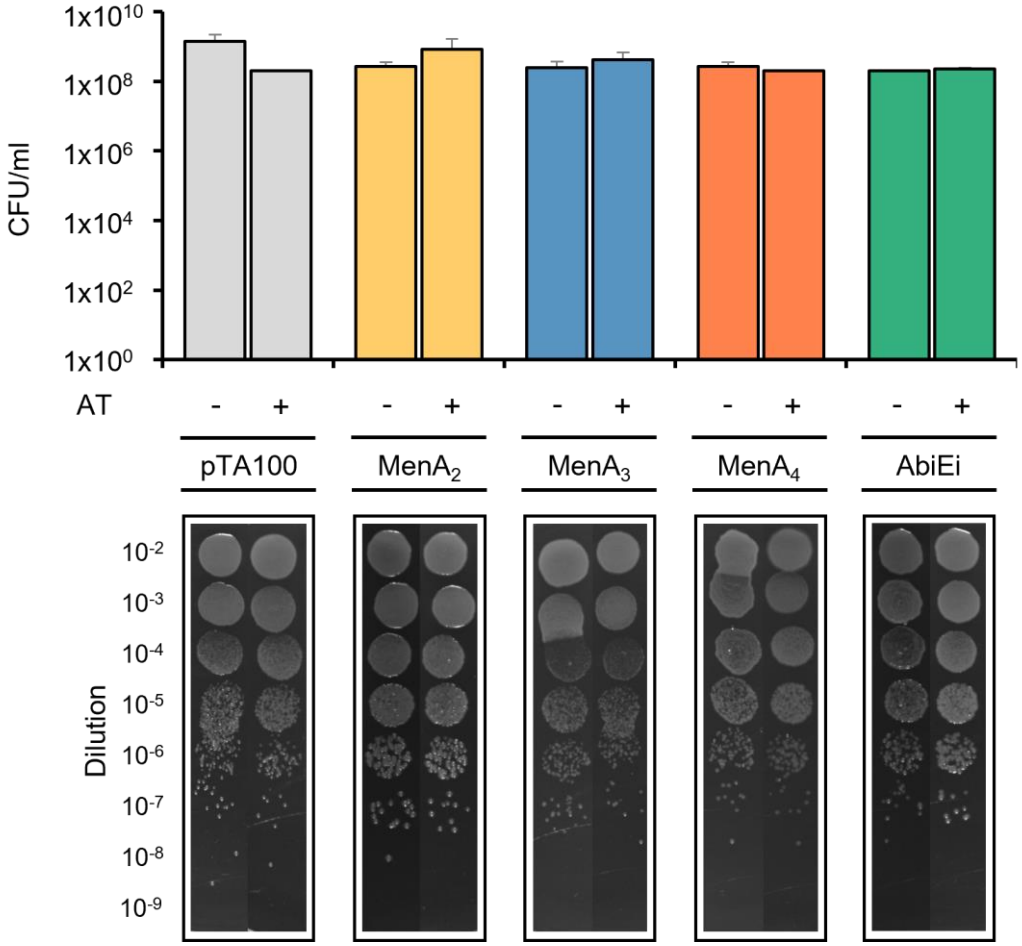


Figure 3.5. MenA antitoxin expression does not affect *E. coli* growth in minimal media. Endpoint viable count toxicity assays of *E. coli* DH5α transformed with the pTA100 vector, pPF660 (*menA₂*), pPF656 (*menA₃*), pPF658 (*menA₄*) or pTRB481 (*abiEi*). Overnight cultures were re-seeded into fresh LB supplemented with Sp and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Sp alone, or with IPTG to induce antitoxin gene expression. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "AT" = antitoxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

In contrast to antitoxin expression, toxin expression had a pronounced impact on bacterial growth (Figure 3.6). Whilst $MenT_2$ again failed to cause any growth defect in minimal media, induction of both $MenT_3$ and $MenT_4$ expression led to an approximately 6-log_{10} drop in CFU/ml (Figure 3.6). Induction of AbiEii also inhibited growth (Figure 3.6). However, the corresponding uninduced sample also caused impaired growth in minimal media, despite the presence of D-glu to repress transcription. Presumably toxicity occurred as a result of leaky transcription from the P_{BAD} promoter, which serves to further highlight the higher relative toxicity of AbiEii in *E. coli*.

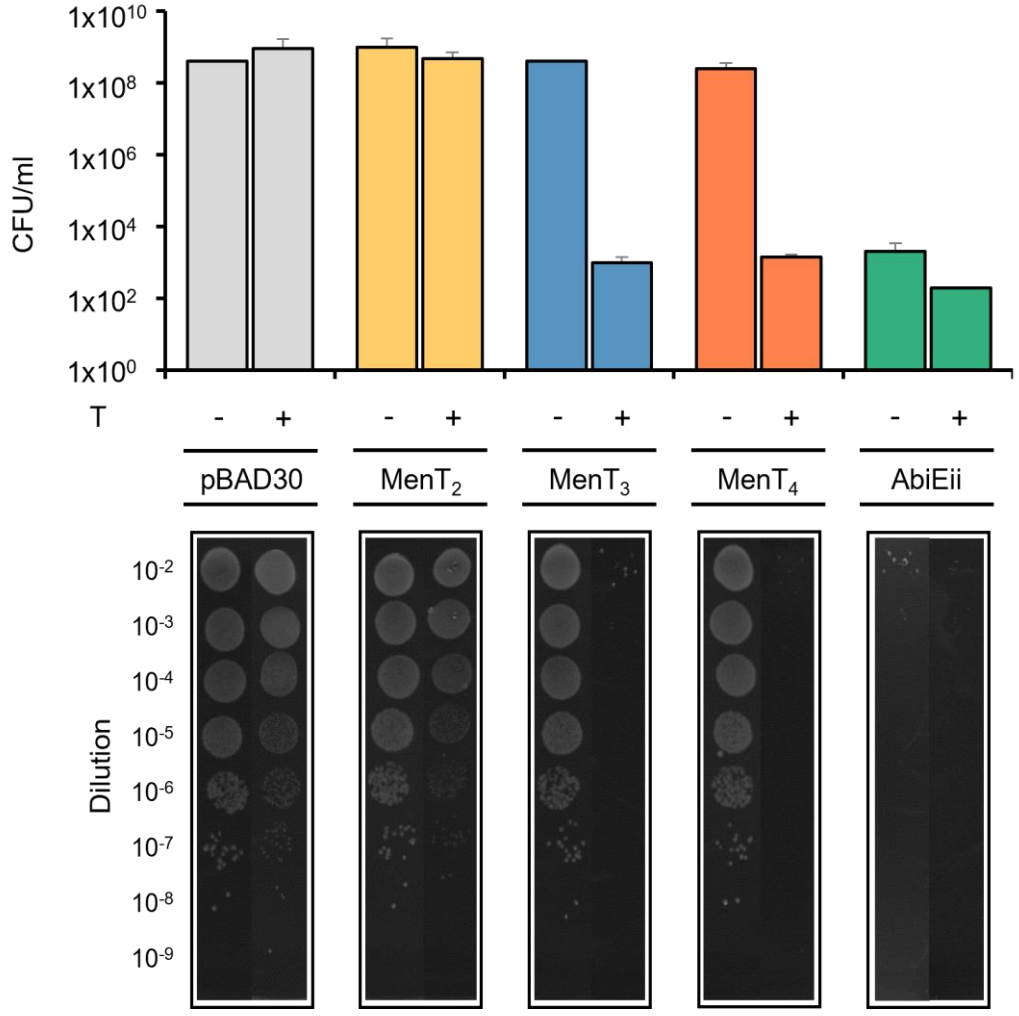


Figure 3.6. $MenT_3$ and $MenT_4$ toxins inhibit *E. coli* growth in minimal media. Endpoint viable count toxicity assays of *E. coli* DH5α transformed with the pBAD30 vector, pTRB480 (*menT₂*), pPF657 (*menT₃*), pPF659 (*menT₄*) or pTRB482 (*abiEii*). Overnight cultures were re-seeded into fresh LB supplemented with Ap and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Ap and either D-glu or L-ara, for repression or induction of gene expression, respectively. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

Having confirmed that two of the three MenT toxins inhibit growth in *E. coli*, all four systems were again tested for TA activity by assessing growth phenotypes of the *E. coli* DH5α dual-plasmid strains under different induction conditions, this time in minimal media (Figure 3.7). Interestingly, the presence of $MenA_3$, regardless of expression state, was again sufficient to nullify $MenT_3$-mediated toxicity (Figure 3.7A). *E. coli* DH5α were therefore co-transformed with pPF657 (*menT₃*) and the pTA100 empty vector and tested by antitoxicity assay as per *E. coli* DH5α pPF656 (*menA₃*) pPF657 (*menT₃*) (Figure 3.7B). Expression of $MenT_3$ reduced growth by >6-log_{10}, including under pTA100 induction conditions. This indicated that the lack of growth inhibition observed for *E. coli* DH5α pPF656 pPF657 was a $MenA_3$-specific effect, demonstrating that $MenA_3$-$MenT_3$ is a TA system, and suggesting that $MenA_3$ is a particularly potent antitoxin.

$MenA_4$-$MenT_4$ behaved slightly differently (Figure 3.7A). Under conditions where antitoxin expression was repressed and toxin expression was induced, the $MenA_4$-$MenT_4$ strain saw a >5-log_{10} reduction in CFU/ml. Co-expression of $MenA_4$ restored viable counts, comparable to uninduced samples, although colony size was noticeably smaller. That $MenA_4$ neutralises $MenT_4$ growth inhibition, albeit with an apparent lingering toxicity, thereby highlights $MenA_4$-$MenT_4$ as a functional TA system in *E. coli*, acting via reversible growth inhibition. Conversely, expressing $MenT_2$ in minimal media made no difference to growth. The consistent lack of toxicity across both nutrient-rich and -limited growth media distinguishes $MenT_2$ from its $MenT_3$ and $MenT_4$ homologues, and shows that the $MenA_2$-$MenT_2$ system does not exhibit TA activity in *E. coli*.

AbiEii again demonstrated a high degree of toxicity compared to the MenT toxins. Samples with both antitoxin and toxin expression repressed saw growth inhibition on par with that of induced AbiEii expression (Figure 3.7A). However, expression of AbiEi restored growth, regardless of whether AbiEii expression was repressed or induced, confirming that AbiEi-AbiEii functions as a TA system in *E. coli*.

Growth curves were also performed with the respective *E. coli* DH5α dual-plasmid strains, using minimal media (MenA-MenT systems) or LB (AbiEi-AbiEii). These assays assessed the kinetic growth dynamics of cognate TA systems under different inducing conditions over a 24 hour period, as determined by regular OD_{600} measurements of bacterial cultures (Figure 3.8).

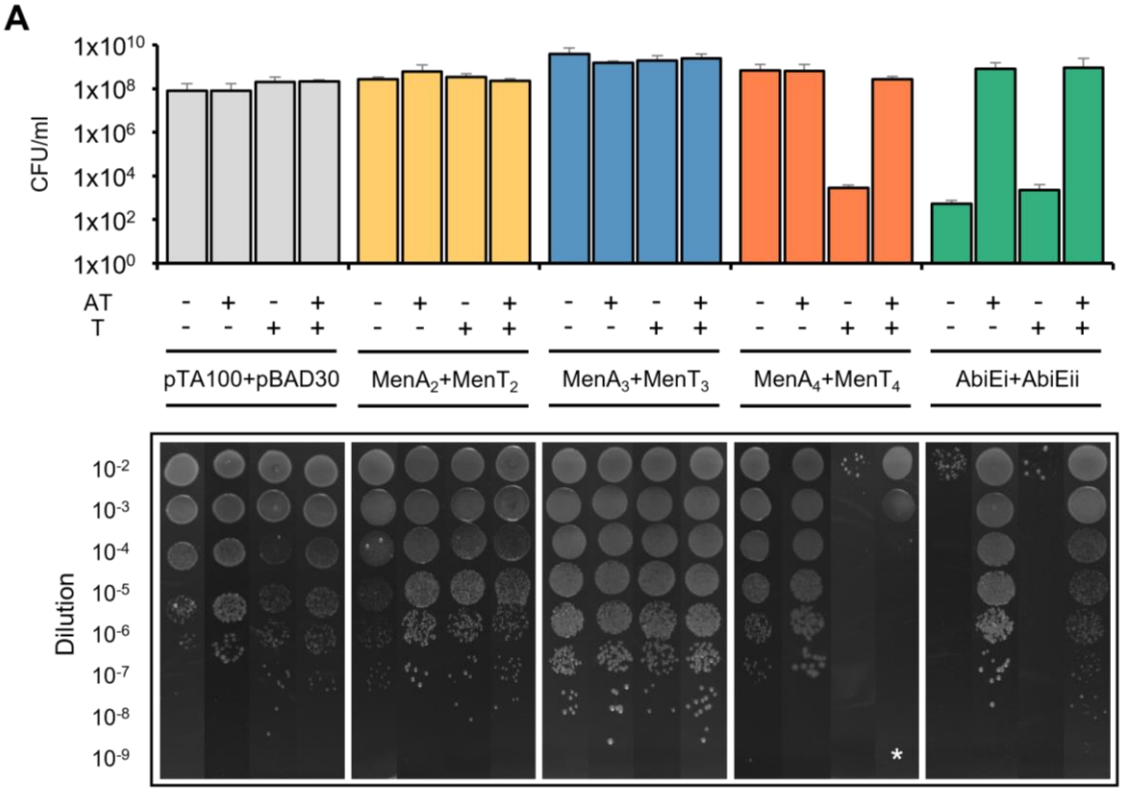


Figure 3.7. MenA$_3$-MenT$_3$ and MenA$_4$-MenT$_4$ function as TA systems in minimal media. (A) Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with either empty vectors or the cognate toxins and antitoxins on separately inducible plasmids (pTA100 + pBAD30, pPF660 (*menA$_2$*) + pTRB480 (*menT$_2$*), pPF656 (*menA$_3$*) + pPF657 (*menT$_3$*), pPF658 (*menA$_4$*) + pPF659 (*menT$_4$*), or pTRB481 (*abiEi*) + pTRB482 (*abiEii*)). Overnight cultures were re-seeded into fresh LB supplemented with Ap, Sp and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Ap and Sp, and with or without D-glu, L-ara and IPTG for repression of toxin expression, induction of toxin expression, and induction of antitoxin expression, respectively. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "AT" = antitoxin, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates). The white asterisk indicates visibly weaker growth, but colony numbers still comparable to the uninduced MenA$_4$+MenT$_4$ sample. **(B)** Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with the pTA100 empty vector and pPF657 (*menT$_3$*), performed as in (A). "V" = vector, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

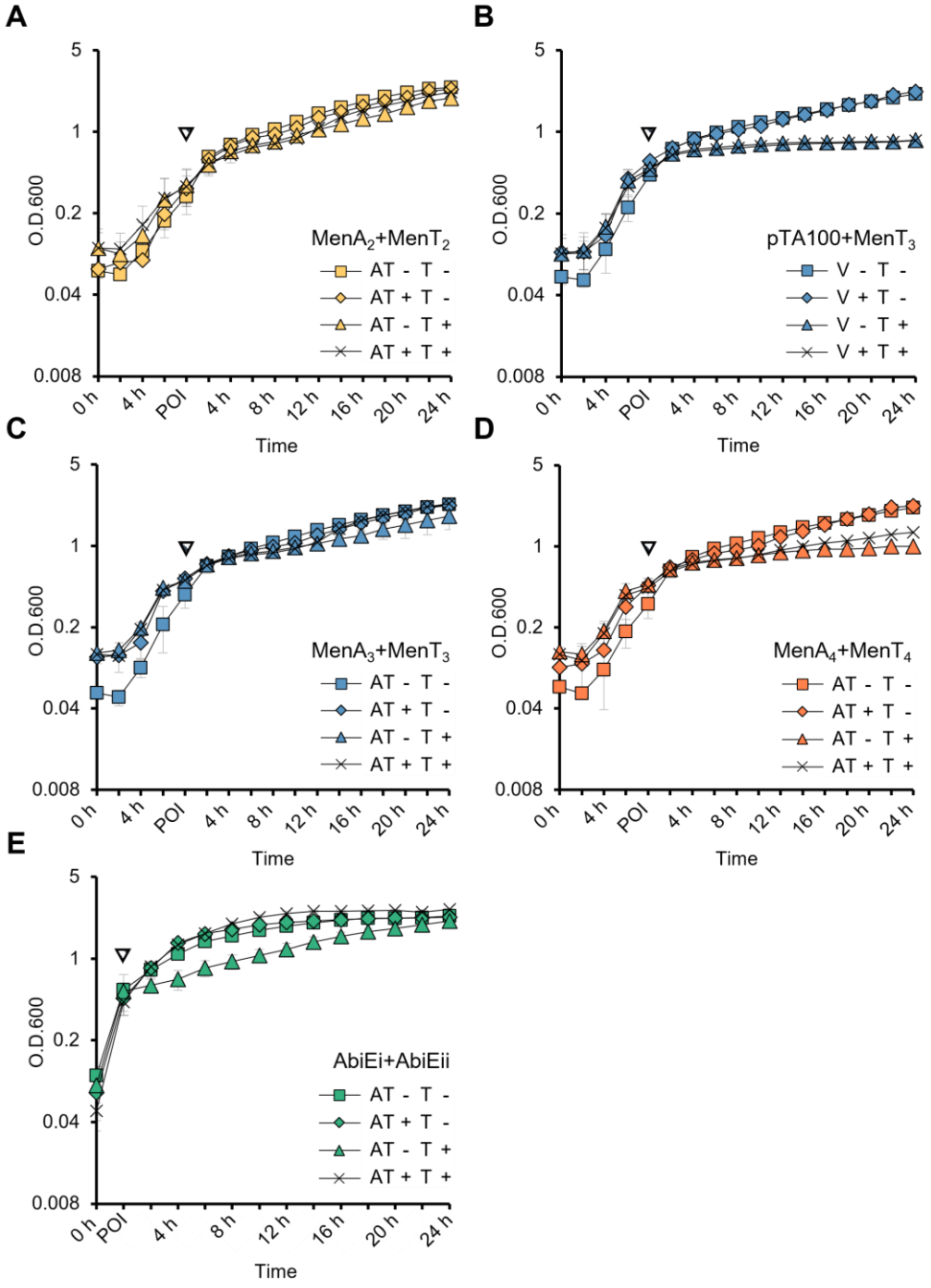


Figure 3.8. MenA$_3$-MenT$_3$ and MenA$_4$-MenT$_4$ have TA activity in *E. coli*. Growth curves of *E. coli* DH5α transformed with cognate antitoxins and toxins on separately inducible plasmids: **(A)** pPF660 (*menA$_2$*) + pTRB480 (*menT$_2$*), M9M; **(B)** pTA100 empty vector + pPF657 (*menT$_3$*), M9M; **(C)** pPF656 (*menA$_3$*) + pPF657 (*menT$_3$*), M9M; **(D)** pPF658 (*menA$_4$*) + pPF659 (*menT$_4$*), M9M; **(E)** pTRB481 (*abiEi*) + pTRB482 (*abiEii*), LB. For all assays, overnight cultures were centrifuged and resuspended in fresh M9M/LB, then reseeded into corresponding wells of a 96-well plate containing M9M/LB, Ap and Sp, and grown to mid-log phase. Samples were then supplemented with either D-glu or L-ara, to repress or induce toxin expression respectively, and with/without IPTG to induce antitoxin expression. Cells were grown for a further 24 h, with OD_{600} measurements taken every 15 min. "AT" = antitoxin, "V" = vector, "T" = toxin; "-" and "+" denote expression state; arrowhead = point of induction (POI). Plotted data represent the mean ± standard deviation (≥ 3 replicates).

$MenT_2$ was again found to be non-toxic in *E. coli* (Figure 3.8A), whereas $MenT_3$ maintained a growth inhibitory effect across the entire experimental time course when expressed alongside the pTA100 empty vector (Figure 3.8B). Consistent with antitoxicity assays (Figure 3.7A), the presence of $MenA_3$, whether induced or not, abolished this toxic phenotype (Figure 3.8C). Induction of $MenT_4$ expression also led to steady growth inhibition (Figure 3.8D). Contrasting with the $MenA_3$-$MenT_3$ antitoxicity phenotype, $MenA_4$ co-expression was unable to fully restore growth to uninduced levels. In the case of AbiEi-AbiEii, co-expression of AbiEi neutralised any effects of AbiEii toxicity (Figure 3.8E). Notably, whilst AbiEii expression alone was initially toxic, cell growth gradually and independently recovered so that, by later time-points, OD_{600} measurements were comparable to those of uninduced samples.

3.5. No interplay between non-cognate DUF1814 toxins and antitoxins

Having shown that $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ function as TA modules in *E. coli*, the next step was to explore potential interplay between these systems. All three MenT toxins are members of the DUF1814 protein family, along with AbiEii, whilst the corresponding MenA and AbiEi antitoxins all belong to the CL0578 superfamily. Considering the homologous nature of these systems, the possibility of cross-talk between non-cognate toxins and antitoxins seemed plausible. Indeed, non-cognate TA systems have been previously shown to exhibit cross-talk: for example, the CcdA antitoxin from *E. chrysanthemi* can inhibit activity of the CcdB toxin homologue from the *E. coli* F plasmid[218]. As a result, *E. coli* DH5α were transformed with combinations of non-cognate toxins and antitoxins and tested for reversible growth inhibition using antitoxicity assays (Figure 3.9). However, no cross-talk was detected between any of the TA pairs.

3.6. Discussion

The functional characterisation of the cognate *menA-menT* gene pairs was carried out to determine whether they act as TA systems, as previously predicted[11]. Both $MenT_3$ and $MenT_4$ exhibited high levels of toxicity in minimal media, which was neutralised by co-expression of the cognate antitoxin (Figure 3.7). Expression of $MenT_2$, however, exerted no toxic effects in LB or minimal media. Therefore, of the three tested MenA-MenT systems, both $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ function as TA systems in *E. coli* through reversible growth inhibition (Figure 3.10), whereas $MenA_2$-$MenT_2$ does not.

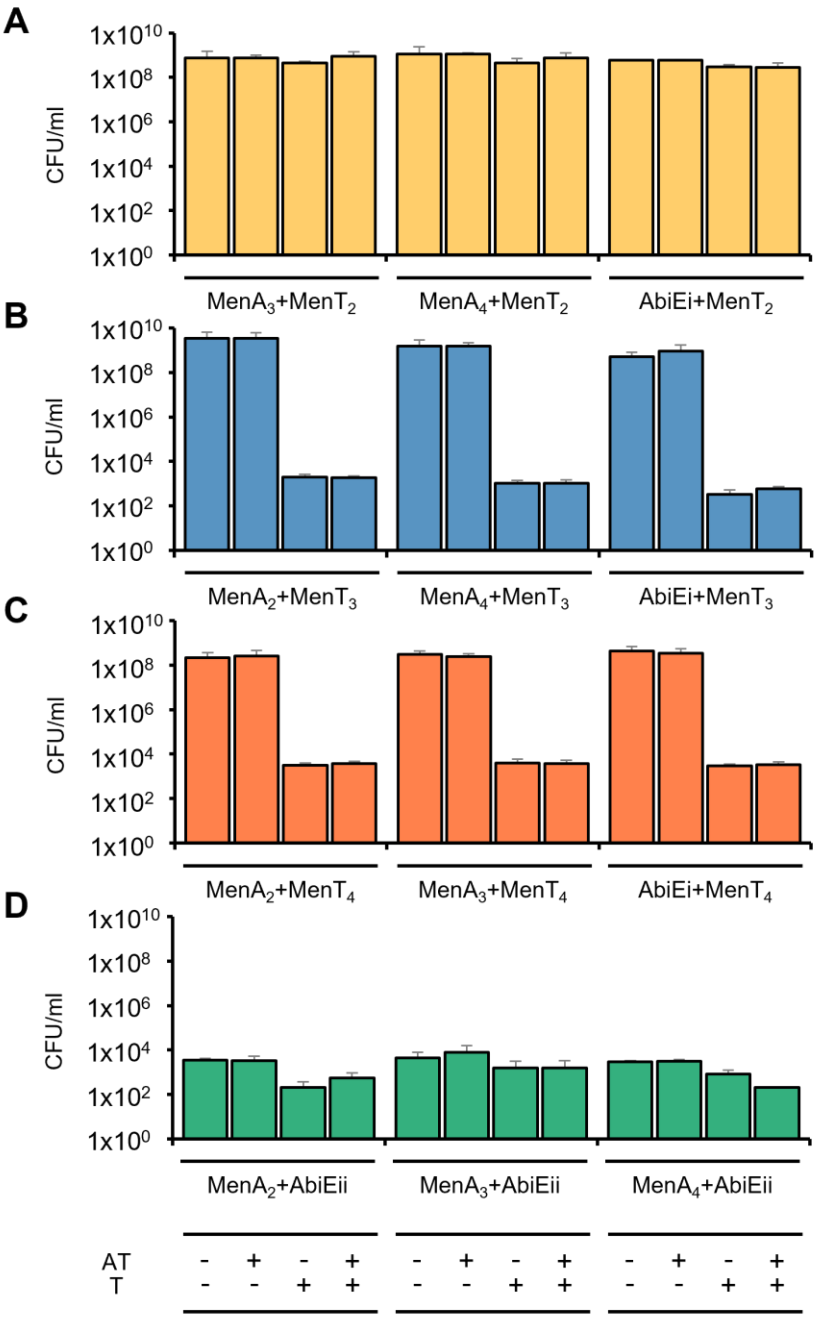


Figure 3.9. DUF1814 toxins and non-cognate antitoxins do not interact. Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with non-cognate antitoxins and toxins on separately inducible plasmids: **(A)** pTRB480 (*menT₂*) and either pPF656 (*menA₃*), pPF658 (*menA₄*) or pTRB481 (*abiEi*); **(B)** pPF657 (*menT₃*) and either pPF660 (*menA₂*), pPF658 (*menA₄*) or pTRB481 (*abiEi*); **(C)** pPF659 (*menT₄*) and either pPF660 (*menA₂*), pPF656 (*menA₃*), or pTRB481 (*abiEi*); **(D)** pTRB482 (*abiEii*) and either pPF660 (*menA₂*), pPF656 (*menA₃*) or pPF658 (*menA₄*). For all assays, overnight cultures were re-seeded into fresh LB supplemented with Ap, Sp and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Ap and Sp, and with or without D-glu, L-ara and IPTG for repression of toxin expression, induction of toxin expression, and induction of antitoxin expression, respectively. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "AT" = antitoxin, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

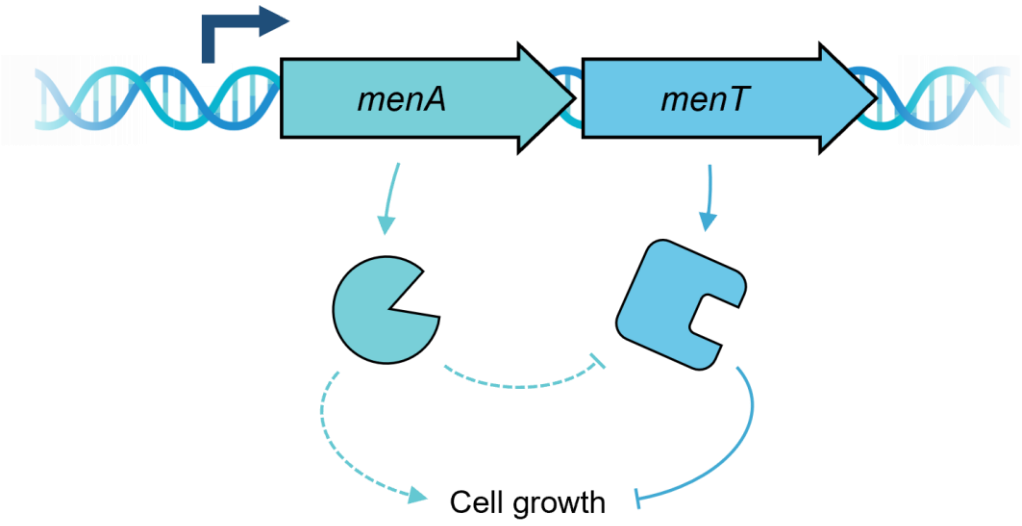


Figure 3.10. Schematic summary of Chapter 3 results: MenA-MenT TA systems reversibly inhibit *E. coli* growth. The $MenT_3$ and $MenT_4$ toxins inhibit *E. coli* cell growth through an unknown mechanism. $MenA_3$ and $MenA_4$ antitoxins neutralise cognate toxin toxicity, however, the mode of antitoxicity is also unknown. Solid lines indicate known activity; dashed lines indicate unknown activity.

All three MenT toxins and AbiEii have been identified as belonging to the NTase-like DUF1814 protein family, and AbiEii has previously been reported to bind GTP specifically, a key indicator of NTase activity[83]. However, the mechanism of $MenT_3$ and $MenT_4$ toxicity could not be determined from these phenotypic growth assays. Despite their shared homology, AbiEii appears considerably more toxic in *E. coli* than $MenT_3$ and $MenT_4$; where the latter two toxins only prevent colony formation in minimal media, AbiEii is toxic in both nutrient-rich LB and nutrient-sparse M9M (Figures 3.3 and 3.6). Therefore, each toxin may have a distinct mode of activity: AbiEii toxicity may simply represent a more potent mechanism of action specific to that toxin; alternatively, the cellular target for AbiEii may potentially be more mechanistically conserved or available in *E. coli* than it is for $MenT_3$ or $MenT_4$. Indeed, *M. tuberculosis* (Actinobacteria), *S. agalactiae* (Firmicutes) and *E. coli* (Proteobacteria) belong to distinct phyla[376], which may reflect the disparate toxicity of the DUF1814 toxins in *E. coli*.

The previously reported upregulation of $menA_2$ and $menT_2$ transcription in *M. tuberculosis* persister cells suggests a potential role for $MenA_2$-$MenT_2$ in controlling growth[251,252], consistent with TA activity. Surprisingly, however, no growth loss phenotype was observed following $MenT_2$ expression. This may be because $MenA_2$-$MenT_2$ is simply not a TA system, and its association with *M. tuberculosis* persisters may be unrelated to TA activity. Alternatively, it may instead be the case that this system does not function effectively in *E. coli*, for example because of improper folding or a lack of expression, or that the $MenT_2$ target

is not conserved in *E. coli*. However, these systems were also characterised in our 2020 study by Cai *et al*., to which the results from this work contributed. Corroborating the results from *E. coli* toxicity assays presented here, our collaborators also showed in the Cai *et al*. study that $MenA_2$-$MenT_2$ fails to function as a TA system in *M. smegmatis*[98]. This was surprising, due to the closer biological relevance of *M. smegmatis* to *M*. *tuberculosis*, and perhaps indicative of a misidentification of $MenA_2$-$MenT_2$ as a putative TA system, despite DUF1814 homology and the classic gene-pair TA architecture. Consequently, whether $MenA_2$-$MenT_2$ functions as a TA system at all remains to be determined. Further work would need to be conducted assessing growth phenotypes in *M*. *tuberculosis* itself, to confirm or reject this predicted function.

In contrast to $MenA_2$-$MenT_2$, the Cai *et al*. study confirmed that $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ both function as TA systems in *M. smegmatis*, indicating a conserved mode of activity[98]. We also demonstrated that $MenT_3$ exhibited the greatest toxicity among the MenT toxins, as uninduced $MenT_3$ toxin expression was still capable of inhibiting growth in *M. smegmatis*, a phenotype not shared by $MenT_4$[98]. This corroborated the results presented above (Figure 3.3), where expression of $MenT_3$ in LB attenuated the growth of bacterial colonies, which was not replicated by $MenT_4$ expression. $MenA_3$-$MenT_3$ was also characterised in *M*. *tuberculosis*, where we found that $MenT_3$ exhibited strong toxicity *in vivo* in *M*. *tuberculosis*, which was rescuable by co-expression of $MenA_3$[98]. We thus demonstrated that $MenA_3$-$MenT_3$ is functionally relevant to *M*. *tuberculosis* growth, although future steps exploring the physiological role of $MenA_3$-$MenT_3$ within the context of *M*. *tuberculosis* growth regulation will need to be undertaken.

A notable observation from growth curves was that samples expressing AbiEii were able to recover growth independently of AbiEi expression (Figure 3.8). By the 24 hour mark, cell growth, as calculated by OD_{600} measurements of bacterial culture, was almost comparable to uninduced samples. This phenotype might result from the acquisition of loss-of-function mutations in the plasmid-encoded *abiEii* gene or P_{BAD} promoter, as has been reported previously for MazF expression in *M. smegmatis*[377], possibly selected for to compensate for AbiEii toxicity.

The mechanism of $MenA_3$ and $MenA_4$ antitoxicity is not discernible from these experiments. It may result from direct toxin inhibition by the antitoxin, as per type II systems, or via antagonistic activity of the toxin and antitoxin against a shared target, as with previously reported type IV TA modules[81,82] (Figure 3.10).

Interestingly, the $MenA_3$ and $MenA_4$ antitoxins produce different growth phenotypes in *E. coli*. $MenA_3$, for example, appears to generate remarkably effective antitoxicity; its mere presence alongside $MenT_3$, even when uninduced, is capable of completely abolishing toxicity (Figures 3.7 and 3.8). While the antitoxic mechanism is unclear from the assays presented above, this might reflect a greater prevalence of, or affinity for, the antitoxin target in *E. coli*, compared to that of $MenT_3$. Alternatively, when expressed in its native environment, the effect of $MenT_3$ toxicity on *M*. *tuberculosis* growth may be significantly more severe, thereby necessitating the robust antitoxicity of $MenA_3$. On the other hand, $MenA_4$ expression appeared to confer comparably less protection against its cognate toxin than $MenA_3$ (Figure 3.8). Moreover, expressing $MenA_4$ alone also appeared to destabilise bacterial growth, not by reducing viable counts, but rather by generating a larger and more irregular colony morphology (Figures 3.2 and 3.5). This observation suggested that $MenA_4$-$MenT_4$ might interact with key components of cell structure, in a manner akin to the CbeA-CbtA *E. coli* type IV system, which antagonistically interacts with the FtsZ and MreB cytoskeletal proteins to reversibly control growth[81,82].

Despite their apparent homology – all four toxins are DUF1814 proteins, and the corresponding antitoxins belong to the CL0578 superfamily – no cross-talk was detected between any of the TA systems tested (Figure 3.9). This may also be due to the distinct phylogenetic lineages of *M*. *tuberculosis* and *S. agalactiae*[376]; while these gene pairs may share a common origin, any conserved features of the DUF1814 toxins and CL0578 antitoxins are likely to have undergone significant evolutionary change, potentially to fulfil distinct biological functions, target specific processes, and have isolated mechanisms of regulation.

4. Autoregulation of MenA-MenT TA systems

4.1. Structural commonality between AbiEi and $MenA_4$ antitoxins suggests a shared mechanism of autoregulation

Having determined that $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ function as TA modules in *E. coli*, the next aim was to explore the regulatory mechanisms governing expression from the TA operon. Autoregulation is a relatively common function among type II TA systems, where the antitoxin binds the operon promoter, often in tandem with the cognate toxin, to repress transcription (See, *1.5*). Though less evidenced among other TA classifications, the AbiEi-AbiEii system from *S. agalactiae* was the first type IV TA module shown to negatively autoregulate its own expression[83]. This activity is important in the context of MenA-MenT TA characterisation, due to the reported homology of AbiEi-AbiEii to the MenA-MenT systems from *M*. *tuberculosis*[11,83].

The COG5340 AbiEi antitoxin contains an uncharacterised bifunctional CTD, essential for both antitoxicity and autoregulation, and a conserved N-terminal wHTH DNA-binding domain[83,84]. AbiEi binds as independent monomers to two 23 base pair (bp) IRs in the *abiEi-abiEii* promoter; the wHTH domain of each protein binds a conserved 11 bp outward-facing region of the IR, whilst the CTD interacts with the remaining 12 bp internal section[83,84]. This interaction is dependent on a net-positive charge which extends across the length of the DNA-facing protein surface[84]. AbiEi autoregulation is positively cooperative; binding of one AbiEi monomer to either IR cooperatively enhances the binding of a second monomer to the remaining free IR[84]. The subsequent occupation of both IRs induces bending of the operator region by 72°; this is proposed to improve AbiEi CTD positioning and IR binding, and may also facilitate AbiEi-AbiEi interactions to further assist cooperativity[84]. Notably, AbiEii does not enhance transcription repression, indicating that AbiEi-AbiEii doesn't follow the conditional cooperativity model common to type II TA modules[83] (See, *1.5.1*).

In addition, the X-ray crystal structure of $MenA_4$, a COG5340 protein and predicted structural homologue of AbiEi[84], was solved in 2009 to 1.93 Å[378]. $MenA_4$ features two structural domains: one, a C-terminal motif which represents a novel structural fold, is proposed to be important for metal binding; while the second domain is an N-terminal wHTH DNA-binding motif[378]. The protein surface is also differentially charged; one face of the CTD is negatively charged and includes the predicted metal-binding site. The opposite side exhibits an electropositive charge stretching the length of the protein, and includes a potential nucleic acid-binding motif in the wHTH N-terminal domain (NTD)[378]. The authors proposed that this positively charged domain mediates $MenA_4$-DNA binding[378], consistent with AbiEi-DNA

interactions[84]. They subsequently demonstrated that $MenA_4$ did indeed bind DNA, and that the addition of DNA oligonucleotides enhanced $MenA_4$ stability non-specifically, preventing spontaneous $MenA_4$ precipitation in solution[378].

In accordance with these reports, Phyre2 modelling was performed on the remaining MenA antitoxins to search for predicted structural homologues[361]. Consistent with $MenA_4$ and AbiEi, searches with $MenA_2$ and $MenA_3$ revealed homology to several N-terminal wHTH DNA-binding proteins. As a result, it was hypothesised that the MenA antitoxins might also share the ability to negatively autoregulate their cognate promoters, in a manner similar to AbiEi autoregulation.

4.2. Construction of TA promoter *lacZ* reporter constructs

To assess promoter activity, the predicted TA promoters were first cloned into pRW50[338], a promoterless *lacZ* reporter plasmid, to generate promoter (P)-*lacZ* fusion constructs. Any innate promoter activity drives expression of the *lacZ* reporter gene and synthesis of β-galactosidase (β-gal) (Figure 4.1A). The amount of β-gal produced is approximately proportional to promoter activity and *lacZ* transcription, and can be measured by the addition of ONPG, a synthetic β-gal substrate. ONPG is cleaved by β-gal to release galactose and o-nitrophenol, the latter of which produces a yellow colour which is then measured fluorometrically. The resulting values are expressed as Miller Units, which represent the relative promoter activity[379].

The 500 bp region immediately upstream of each *menA* translational start site, predicted to contain the cognate TA promoter, was amplified by PCR from *M. tuberculosis* H37Rv genomic DNA. The *menA₂-menT₂* upstream region was cloned as an EcoRI/EcoRI fragment into EcoRI-digested pRW50, producing pTRB485 (P_{menA_2}). The upstream regions for *menA₃-menT₃* and *menA₄-menT₄* were cloned as EcoRI/HindIII fragments into EcoRI/HindIII-digested pRW50, generating pTRB483 (P_{menA_3}) and pTRB484 (P_{menA_4}), respectively.

The *abiEi-abiEii* promoter (P_{abiEi}) was also tested, as a positive control of promoter activity. P_{abiEi} had previously been isolated to a 100 bp region upstream of the *abiEi* gene and shown to exhibit strong levels of promoter activity[83]. The P_{abiEi}-*lacZ* fusion was constructed by cloning the 99 bp upstream of *abiEi-abiEii*, amplified from pPF680 as an EcoRI/HindIII fragment, into EcoRI/HindIII-digested pRW50, creating pTRB486.

4.3. MenA$_4$ negatively autoregulates MenA$_4$-MenT$_4$ expression

E. coli DH5α were separately transformed with either pRW50, as a vector-only control, or one of the four TA P-*lacZ* fusions, and assessed for promoter activity (Figure 4.1B). Assays were performed as described in Materials and Methods, *2.5*. Both the *menA*$_2$ and *menA*$_3$ promoters exhibited minimal to no levels of activity relative to the vector-only control. In contrast, both P*menA*$_4$ and P*abiEi* reporters demonstrated *lacZ* expression, though to varying degrees, with P*abiEi* considerably more active than P*menA*$_4$.

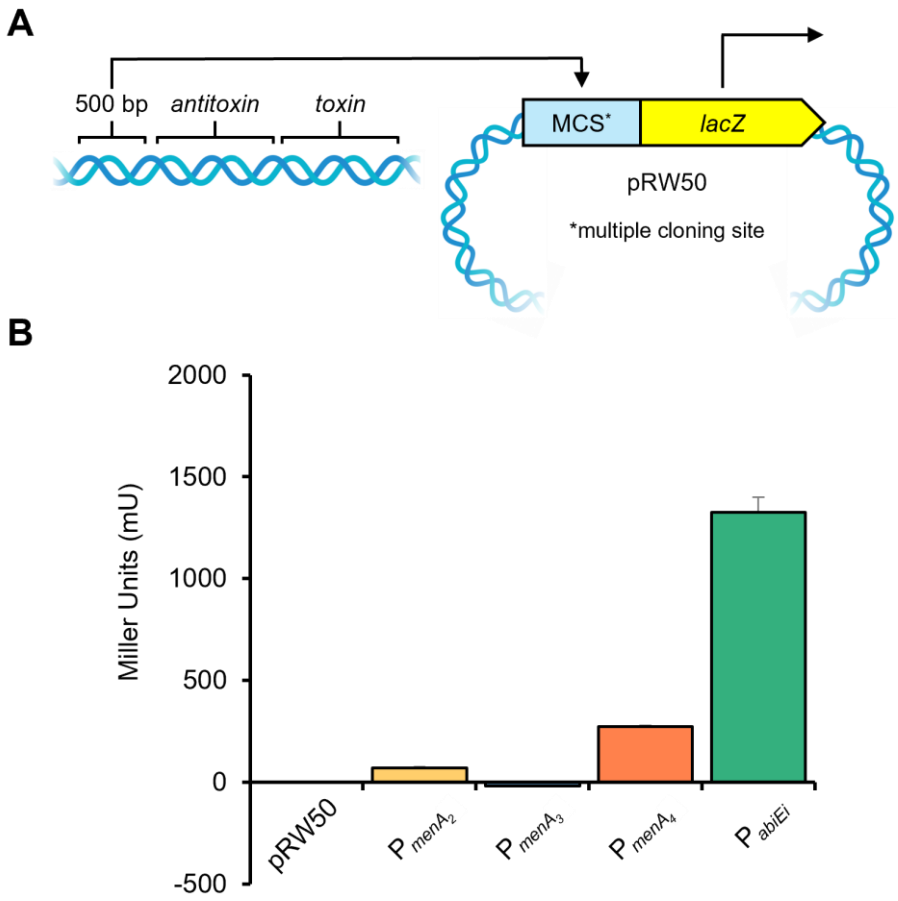


Figure 4.1. Putative *M. tuberculosis* TA promoters display varying activity in *E. coli*. (A) Schematic diagram of the generation of *lacZ* reporter constructs. The 500 bp upstream genomic region of *menA*$_2$-*menT*$_2$, *menA*$_3$-*menT*$_3$ and *menA*$_4$-*menT*$_4$, and the 99 bp upstream region of *abiEi*-*abiEii*, were cloned into the pRW50 promoterless *lacZ* vector. Any promoter activity from the cloned genomic regions drives expression of the *lacZ* gene. **(B)** Promoter activity assays of *E. coli* DH5α transformed with either the pRW50 vector, pTRB485 (P*menA*$_2$), pTRB483 (P*menA*$_3$), pTRB484 (P*menA*$_4$), or pTRB486 (P*abiEi*). Detection of promoter activity was performed as described in Materials and Methods, *2.5*. Plotted data are normalised to the pRW50 vector-only control and represent the mean ± standard deviation (≥ 3 replicates).

As mentioned previously, the type IV AbiEi antitoxin autoregulates the TA operon via transcriptional repression of the cognate promoter[83,84]. To determine whether $MenA_4$ similarly represses promoter activity, the two active P_{menA_4} and P_{abiEi} reporter constructs were paired with IPTG-inducible plasmids expressing the cognate antitoxins (pPF658 and pTRB481, respectively), which were then used to co-transform *E. coli* DH5α. Any negative autoregulation resulting from antitoxin expression would be detected by a quantifiable drop in β-gal production and ONPG cleavage compared to the uninduced sample (Figure 4.2A).

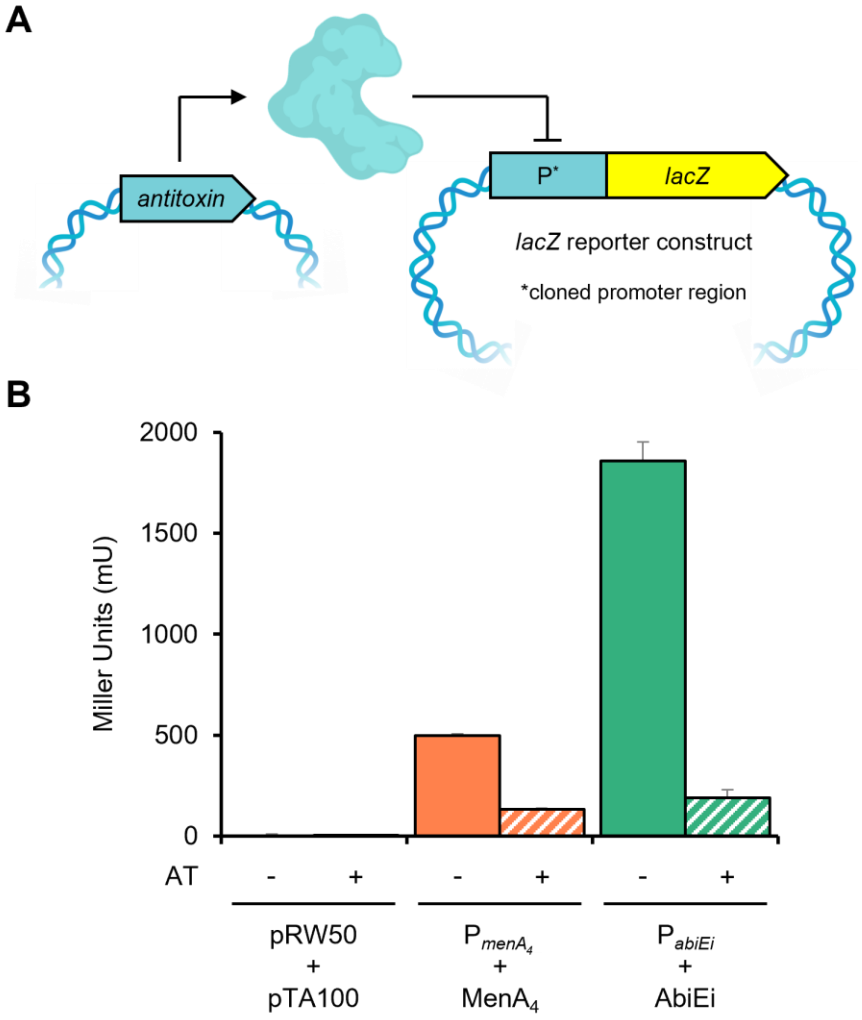


Figure 4.2. $MenA_4$-$MenT_4$ is a negatively autoregulating TA system in *E. coli*. (A) Schematic diagram of antitoxin-mediated autoregulation of P-*lacZ* fusion activity. The antitoxin, expressed from an IPTG-inducible plasmid, suppresses promoter activity from the *lacZ* reporter construct, inhibiting expression of the *lacZ* gene. **(B)** Antitoxin autoregulation assays testing identified active promoters. *E. coli* DH5α were transformed with the pRW50 + pTA100 empty vectors, pTRB484 (P_{menA_4}) + pPF658 (*menA*$_4$), or pTRB486 (P_{abiEi}) + pTRB481 (*abiEi*). Promoter activity was measured with or without induction of the cognate antitoxin (AT, ± IPTG), as described in Materials and Methods, *2.5*. Plotted data are normalised to the uninduced vector-only control and represent the mean ± standard deviation (≥ 3 replicates).

These strains were tested for promoter activity as per Materials and Methods, *2.5*, alongside an *E. coli* DH5α pRW50 pTA100 control strain (Figure 4.2B). Both $MenA_4$ and AbiEi expression resulted in reduced promoter activity relative to the corresponding uninduced control. This confirmed previous reports that AbiEi negatively autoregulates its cognate promoter, and also shows that $MenA_4$ shares this autoregulatory mechanism (Figure 4.2B).

4.4. Discussion

The involvement of MenA antitoxins in transcriptional autorepression was assessed by promoter activity assays, alongside the AbiEi antitoxin from *S. agalactiae*. Confirming previous reports, AbiEi expression negatively autoregulated the cognate promoter, causing a robust drop in β-gal production from the P_{abiEi}-*lacZ* fusion (Figure 4.2B). Similarly, $MenA_4$ also autoregulated its cognate promoter, with $MenA_4$ expression resulting in a >70% fall in promoter activity (Figures 4.2B and 4.3).

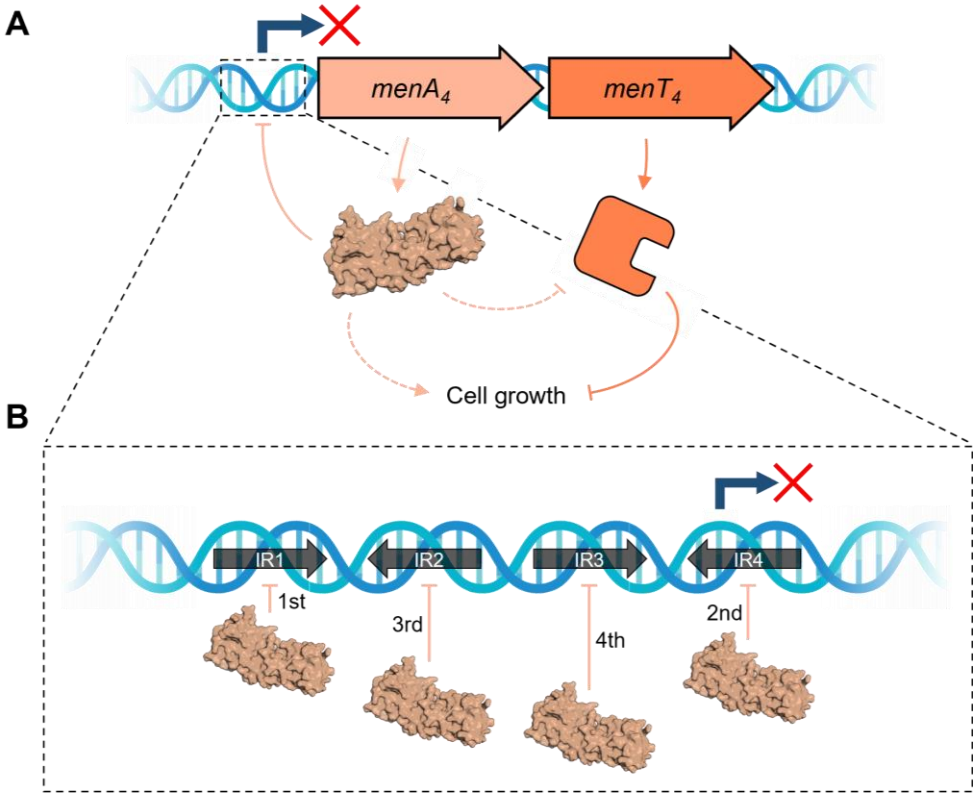


Figure 4.3. Schematic summary of Chapter 4 results: $MenA_4$ negatively autoregulates the cognate TA operon. (A) In addition to neutralising $MenT_4$ toxicity, the $MenA_4$ antitoxin (PDB: 1ZEL) represses transcription from the *menA*$_4$-*menT*$_4$ promoter in *E. coli*. **(B)** The promoter architecture and sequential binding affinity of $MenA_4$ for promoter IR DNA was revealed in a separate study by Beck *et al.*, 2020[380].

These findings were expanded on in a 2020 study by Beck *et al*., to which the experiments presented in Figures 4.1 and 4.2 contributed[380]. In this study, we solved the AbiEi X-ray crystal structure to 1.83 Å, and demonstrated that AbiEi shares a conserved NTD, CTD, and overall charge features with $MenA_4$[380]. However, despite the structural homology between the AbiEi-AbiEii and $MenA_4$-$MenT_4$ systems, the respective promoters and DNA-binding mechanisms have evolved distinct differences[380]. For example, although the 23 bp IR sequences of P_{abiEi} and P_{menA_4} are conserved, P_{menA_4} contains an additional IR pair[380]. Moreover, where AbiEi binds IRs with positive cooperativity, $MenA_4$ instead acts through negative cooperativity, interacting with the four cognate IRs via sequential binding affinities[380] (Figure 4.3B).

It was earlier described that no antitoxic interplay was observed between non-cognate MenA antitoxins and MenT toxins during growth assays in *E. coli* (See, *3.5*; Figure 3.9). Whilst antitoxicity assays did not investigate the potential for cross-reactive autoregulation, the Beck *et al*. study went a step further, confirming that neither $MenA_4$ nor AbiEi could bind the non-cognate P_{abiEi} or P_{menA_4} promoter *in vitro*[380]. The distinct promoter architectures and sequences of P_{menA_4} and P_{abiEi} provide a rationale for these observations. However, it would be useful to further explore the potential for regulatory interplay between all MenA-MenT systems, by performing additional promoter activity assays testing the MenA antitoxins for inhibitory activity against non-cognate promoters. While the mechanisms of autoregulation may differ sufficiently to prevent cross-regulation, the consequences of antitoxin-DNA binding are likely comparable between systems. For $MenA_4$, it was suggested that interactions with the P_{menA_4} IRs sterically obstruct σ factor binding to prevent transcription by RNA polymerase, thereby resulting in autoregulation of the cognate operon[380].

In contrast to the results from P_{abiEi}-*lacZ* and P_{menA_4}-*lacZ* reporters, no β-gal activity was detected for either the $menA_2$ or $menA_3$ putative promoters tested (Figure 4.1B). The lack of activity for P_{menA_3} was intriguing as, despite both $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ functioning as TA systems in *E. coli*, only P_{menA_4} exhibited activity in *lacZ* reporter assays. On the other hand, the minimal activity exhibited by P_{menA_2} is perhaps unsurprising: $MenA_2$-$MenT_2$ does not appear active in *E. coli*, and therefore the prediction that a $menA_2$-$menT_2$ promoter lies in the 500 bp upstream region of these uncharacterised genes may be flawed. Alternatively, it may be a case that an insufficiently large upstream region was chosen for P_{menA_2} and P_{menA_3} analysis, and that testing in *E. coli* is insufficient to thoroughly assess *M. tuberculosis* promoter activity. Indeed, compared to *E. coli,* mycobacterial promoters are often more complex, potentially extending up to 2 kb from the transcriptional start site,

missing canonical features of operon architecture such as the -35 element, and generating a transcriptome where 25% of transcripts are leaderless[381–383]. Moreover, the wealth of σ factors and environment-responsive transcription factors present in *M*. *tuberculosis* encourages greater promoter sequence variation, and contributes further complexity to transcriptional regulation[384,385].

Recent studies have, however, provided conflicting evidence regarding $MenA_3$ and P_{menA_3} autoregulation. Consistent with the findings described here, the Beck *et al*. study also noted that $MenA_3$ was unable to bind its putative promoter via *in vitro* DNA-binding assays[380]. Interestingly, we instead found that $MenA_3$ unexpectedly bound the non-cognate IRs of P_{abiEi}, but not P_{menA_4}[380]. Whilst this binding was weak and had no demonstrable cooperativity, this result suggested that $MenA_3$ was indeed capable of binding DNA, but we had not found the correct region[380]. Genomic analyses found that P_{abiEi} aligns with multiple locations in the *M*. *tuberculosis* genome[380]. Therefore, it was suggested that $MenA_3$ might interact with other gene promoters, functioning as a general or condition-specific transcription regulator[380], as has been described for other TA systems[88,152–154]. However, contrary to our results, Yu *et al*., 2020, separately reported that the 500 bp region upstream of *menA₃* was in fact active in *lacZ* reporter assays performed from a chromosomal reporter in *M. smegmatis*[99]. Moreover, they also found that $MenA_3$ was sufficient to repress promoter activity, with or without expression[99]. The *in vitro* DNA-binding assays performed in the Beck *et al*. paper only tested a 131 bp region upstream of the *menA₃* translational start site, consistent with the location of the P_{menA_4} IRs[380]. In doing so, we likely overlooked relevant promoter architecture located further upstream, potentially with functional relevance *in vivo* in *M. smegmatis* but not in *E. coli*, as suggested by the absence of P_{menA_3}-*lacZ* reporter activity presented here (Figure 4.1B).

Needless to say, future work will need to definitively identify and characterise P_{menA_3}, as well as confirm the absence or presence of $MenA_3$ autoregulation *in vitro* and *in vivo*. In addition, characterisation of P_{menA_4}, as well as the *menA₄-menT₄* operon, can still benefit from further investigation despite the results reported here. Confirmatory reporter assays will be useful to determine promoter activity and autoregulation *in vivo* in *M. smegmatis*. This could be particularly important given the contrasting reports of P_{menA_3} activity. The importance of residues key to $MenA_4$-DNA binding, as highlighted by Beck *et al*.[380], could be confirmed by mutational studies of $MenA_4$ in *lacZ* reporter assays. Similarly, testing mutated P_{menA_4} IR sequences as *lacZ* fusion constructs might further determine their relative importance to autoregulation[380]. Additional mutational studies might also analyse various $MenA_4$ truncations to determine their essentiality to autoregulation, as has been previously done

for AbiEi[83]. Finally, future work should determine the effects of toxin co-expression on autoregulation, and whether it enhances the regulatory effect seen by $MenA_4$ alone. This would further characterise the autoregulation mechanism and inform the TA classification to which $MenA_4$-$MenT_4$ belongs. For example, the observation that AbiEii did not enhance transcriptional repression helped reveal that AbiEi-AbiEii is a type IV TA module[83].

5. Structural characterisation of MenT toxins

5.1. Identification of an additional MenA-MenT family member: $MenA_1$-$MenT_1$

Midway through the project, a collaboration was developed with the Genevaux lab, who happened to also be characterising the MenA-MenT family TA systems. In hindsight this was unsurprising; their group produced one of the two 2014 papers which identified the MenA-MenT family as homologues of putative type IV TA systems (alongside the earlier study by Dy *et al*., 2014)[11,83]. Notably, the Genevaux group revealed the existence of a fourth MenA-MenT family member, Rv0078B (antitoxin)-Rv0078A (toxin), herein referred to as $MenA_1$-$MenT_1$ (personal communication) (Figure 5.1).

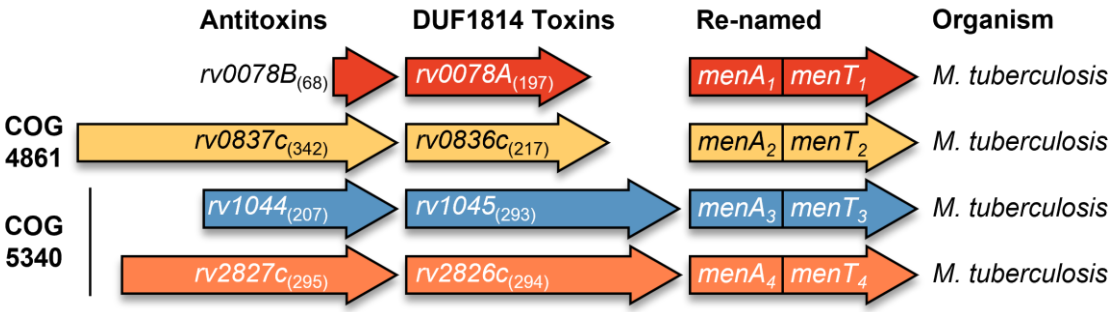


Figure 5.1. *M. tuberculosis* encodes four TA systems with DUF1814 NTase-like toxins. Graphical representation of the four *M. tuberculosis* MenA-MenT TA systems containing DUF1814 NTase-like toxin genes, scaled according to protein AA composition (detailed in parentheses). Where applicable, the assigned COG group of each antitoxin is denoted to the left. Each TA system has been allocated a specific colour, retained throughout the manuscript. Revised nomenclature and the host organism are depicted to the right.

$MenT_1$ shares homology to the other MenT toxins via the conserved NTase-like DUF1814 domain, though with approximately two thirds of the AA content of $MenT_3$ and $MenT_4$, this might reflect a distinct mode of action. The *menT₁* gene is paired with the short upstream *menA₁* gene, which encodes a small 68-AA protein (Figure 5.1). $MenA_1$ is particularly interesting as it has no assigned COG classification, no apparent homology to the other MenA antitoxins (highest AA sequence identity = 18.3%, $MenA_1$ to $MenA_3$; EMBOSS Stretcher), and was flagged by the TASmania database as a putative toxin, with similarity closest to the type I SymE RNase toxin[386]. In spite of this, $MenA_1$-$MenT_1$ was shown to act as a TA system in *M. smegmatis*, where expression of $MenT_1$ inhibited bacterial growth and co-expression of $MenA_1$ rescued growth[386]. This is intriguing and suggests an evolutionary divergence among MenA-MenT TA systems, perhaps to fulfil separate biological functions with antitoxicity controlled through distinct mechanisms.

Having explored *menA-menT* promoter activity and MenA antitoxin autoregulation, and determined that the $MenT_3$ and $MenT_4$ toxins are functional in *E. coli*, the next objective was to characterise the molecular mechanisms of MenT toxicity. To begin investigations, X-ray crystallographic studies were carried out to determine the $MenT_3$ and $MenT_4$ toxin crystal structures. It was predicted that solving the MenT structures would provide detailed insight into their biological roles, utilising *in silico* methods such as structural homology searches and protein conservation databases.

5.2. Construction of MenA/MenT expression plasmids for crystallographic studies

To begin the structural characterisation of the MenA-MenT TA modules, expression constructs first needed to be generated. The *menT₂*, *menT₃* and *menT₄* toxin genes were amplified by PCR from pTRB480, pPF657 and pPF659, respectively, and cloned into pSAT1-LIC via LIC (Materials and Methods, *2.3.2.4*). This resulted in pTRB496 (*menT₂*), pTRB492 (*menT₃*) and pTRB494 (*menT₄*). The pSAT1-LIC plasmid encodes a LIC site that fuses a SENP-cleavable N-terminal His_6-SUMO tag to the target gene. Following the expression and initial purification of the recombinant protein, the His_6-SUMO tag can be cleaved by the addition of hSENP2 protease, leaving untagged, isolated protein (Figure 2.2).

Early trial expressions of pTRB496, pTRB492 and pTRB494 in *E. coli* ER2566 yielded minimal-to-no quantifiable expression (data not shown). Toxin genes were therefore re-cloned into pBAD30 derivatives, as the pBAD backbone had previously proven sufficient for controllable toxin expression during phenotypic testing (See, *3.4*). The $MenA_2$-$MenT_2$ system was herein omitted from further analysis, due to the lack of toxin expression and absence of a functional phenotype in *E. coli* (See, *3.4*). For *menT₃* and *menT₄*, overlap PCRs were performed to fuse the N-terminal His_6-SUMO tag amplified from pSAT1-LIC to each of the target genes. The resulting PCR products were cloned as either KpnI/HindIII fragments into KpnI/HindIII-digested pBAD30 (*menT₃*), or as XmaI/HindIII fragments into XmaI/HindIII-digested pBAD30 (*menT₄*), producing pTRB517 and pTRB544 respectively. To speed up subsequent toxin cloning, a pBAD-based variant of pSAT1-LIC (pTRB550) was developed to facilitate direct LIC cloning of toxin genes into a more compatible His_6-SUMO fusion vector (Materials and Methods, *2.3.2.4*).

5.3. $MenT_3$ crystal structure

To begin with, *E. coli* ER2566 were co-transformed with pRARE, a rare tRNA-encoding expression plasmid (Novagen), pPF656, encoding *menA*$_3$ under the control of an IPTG-inducible promoter, and pTRB517. Proteins were expressed and purified as per Materials and Methods, *2.6* and *2.7*. The final sample, containing isolated and purified $MenT_3$ toxin protein, was used to set and optimise crystal screens; the ensuing crystals were extracted, and X-ray crystallographic data were collected as described in Materials and Methods, *2.8* and *2.9*. The resulting dataset was collected to 1.59 Å but proved insufficient for structure resolution by MR. This was likely due to the low sequence identity of the search models used, chosen based on predicted structural homology by Phyre2, as well as the poor fit of a Phyre2-generated $MenT_3$ structural model[361]. The dataset also proved unsuitable for solving the $MenT_3$ structure *ab initio* using ARCIMBOLDO[351], presumably because of an insufficiently high resolution. As a result, SeMet-derivatised $MenT_3$ protein was expressed and purified from the same expression strain, crystal trays were set using previously optimised $MenT_3$ crystal conditions, and X-ray crystallographic data were collected and used to solve the $MenT_3$ crystal structure by MAD (Materials and Methods, *2.6.1*, *2.7*, *2.8* and *2.9*) (Table 5.1 and Figure 5.2).

The $MenT_3$ crystal structure showed a monomeric, bi-lobed globular protein, with the NTD and CTD connected by a short linker (Figure 5.2A). Surface electrostatics of the $MenT_3$ structure show an electropositive groove leading to a recessed electropositive cavity (Figure 5.2B). This potentially indicates the position of the $MenT_3$ active site; the electropositive surface groove may facilitate interactions with electronegative substrates such as nucleic acids, consistent with its classification as an NTase-like DUF1814 protein and the reported GTP-binding activity of homologous AbiEii[83]. Interestingly, the resolved structure showed density for a phosphoserine (SEP) at S78 (Figure 5.2C), suggesting that at some point during $MenT_3$ expression, a phosphorylation event had occurred. This structure was therefore named $MenT_3$ SEP and is referred to as such herein. The $MenT_3$ SEP structure was next analysed with PDBsum[387], which generated a protein topology schematic showing a mixed helical and anti-parallel β-sheet NTD (AAs 4-150), linked to a distinct CTD helical bundle comprising helices α6 to α11 (Figure 5.2D). The large cavity formed between the $MenT_3$ SEP NTD and CTD corresponds to the predicted electropositive active site (Figure 5.2A and B).

To troubleshoot why the $MenT_3$ SEP S78 was phosphorylated, $MenT_3$ was once again expressed and purified for use in X-ray crystallographic studies. As the original $MenT_3$ protein

sample used to determine the $MenT_3$ SEP crystal structure was co-expressed alongside $MenA_3$, the presence and action of $MenA_3$ was proposed as a possible reason contributing to $MenT_3$ S78 phosphorylation. Therefore, *E. coli* ER2566 was co-transformed with just pRARE and pTRB517, then grown and expressed as described previously (Materials and Methods, *2.6*). Consistent with earlier toxicity and antitoxicity assays, expressing $MenT_3$ in the absence of $MenA_3$ had no noticeable impact on bacterial growth in nutrient-rich media. The resulting $MenT_3$ protein was purified and used to set crystal trays with pre-optimised conditions; crystals were then harvested and X-ray crystallographic data collected as per Materials and Methods, *2.7*, *2.8* and *2.9*. The collected data were then used to solve the $MenT_3$ structure by MR using the $MenT_3$ SEP crystal structure as a search model (Materials and Methods, *2.9*).

The $MenT_3$ crystal structure was resolved to 1.78 Å and confirmed the absence of a phosphoserine at S78 (Table 5.1 and Figure 5.3). The overall structure of $MenT_3$ is almost identical to $MenT_3$ SEP, featuring the distinct NTD and CTD architecture (Figure 5.3A), and electropositive surface groove (Figure 5.3B); aligning the two structures by sequence gave a root mean square deviation (RMSD) of 0.158 Å, between 1808 atoms (Figure 5.3C).

Table 5.1. Crystallographic data collection and refinement statistics

	$MenT_3$ SEP Native	$MenT_3$ SEP Se-Peak	$MenT_3$ SEP Se-High Remote	$MenT_3$ SEP Se-Inflection	$MenT_3$ Native
Data Collection					
PDB ID Code	6Y5U	-	-	-	-
Beamline	Diamond I04	Diamond I03	Diamond I03	Diamond I03	Diamond I24
Wavelength, Å	0.9795	0.9793	0.9641	0.9795	0.9795
Resolution range, Å	47.70 – 1.59 (1.65 – 1.59)[a]	47.78 – 2.19 (2.26 – 2.19)	47.83 – 2.05 (2.11 – 2.05)	53.13 – 2.04 (2.11 – 2.04)	41.25 – 1.78 (1.85 – 1.78)
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Unit cell, *a b c* (Å), α β γ (°)	95.4 95.4 69.0, 90.0 90.0 120.0	95.6 95.6 69.2, 90.0 90.0 120.0	95.7 95.7 69.3, 90.0 90.0 120.0	95.6 95.6 69.3, 90.0 90.0 120.0	95.3 95.3 69.0, 90.0 90.0 120.0
Total reflections	98016 (9668)	36407 (3179)	44514 (3476)	47255 (4637)	692864 (68859)
Unique reflections	49008 (4834)	19130 (1646)	23313 (1788)	23628 (2319)	34876 (3443)

Statistics for the highest-resolution shell are shown in parentheses

Table 5.1 continued. Crystallographic data collection and refinement statistics

	$MenT_3$ SEP Native	$MenT_3$ SEP Se-Peak	$MenT_3$ SEP Se-High Remote	$MenT_3$ SEP Se-Inflection	$MenT_3$ Native
Data Collection					
Multiplicity	2.0	1.9	1.9	2.0	19.9 (20.0)
Completeness (%)	99.95 (99.83)	100.00 (100.00)	100.00 (99.80)	100.00 (99.70)	97.35 (75.14)
Mean I/ σ (I)	16.7	6.5	7.6	8.9	12.64
R_{merge}	0.016 (0.486)	0.055 (0.373)	0.055 (0.522)	0.048 (0.463)	0.097 (5.678)
R_{meas}	0.022 (0.687)	0.077 (0.528)	0.078 (0.739)	0.068 (0.654)	0.099 (5.822)
$CC_{1/2}$	1.0 (0.672)	0.995 (0.803)	0.997 (0.544)	0.997 (0.641)	0.999 (0.587)
Refinement					
R_{work}	0.2024 (0.2924)	-	-	-	0.2065 (0.4160)
R_{free}	0.2242 (0.3108)	-	-	-	0.2227 (0.4384)
No. of non-hydrogen atoms	2494	-	-	-	2288
macromolecules	2213	-	-	-	2209
solvent	281	-	-	-	79
Protein residues	288	-	-	-	288
RMSD (bonds, Å)	0.006	-	-	-	0.011
RMSD (angles, °)	0.940	-	-	-	1.46
Ramachandran favored (%)	97.53	-	-	-	97.90
Ramachandran allowed (%)	2.47	-	-	-	2.10
Ramachandran outliers (%)	0.00	-	-	-	0.00
Average B-factor	34.1	-	-	-	60.28
macromolecules	33.4	-	-	-	60.54
solvent	39.4	-	-	-	53.22

Statistics for the highest-resolution shell are shown in parentheses

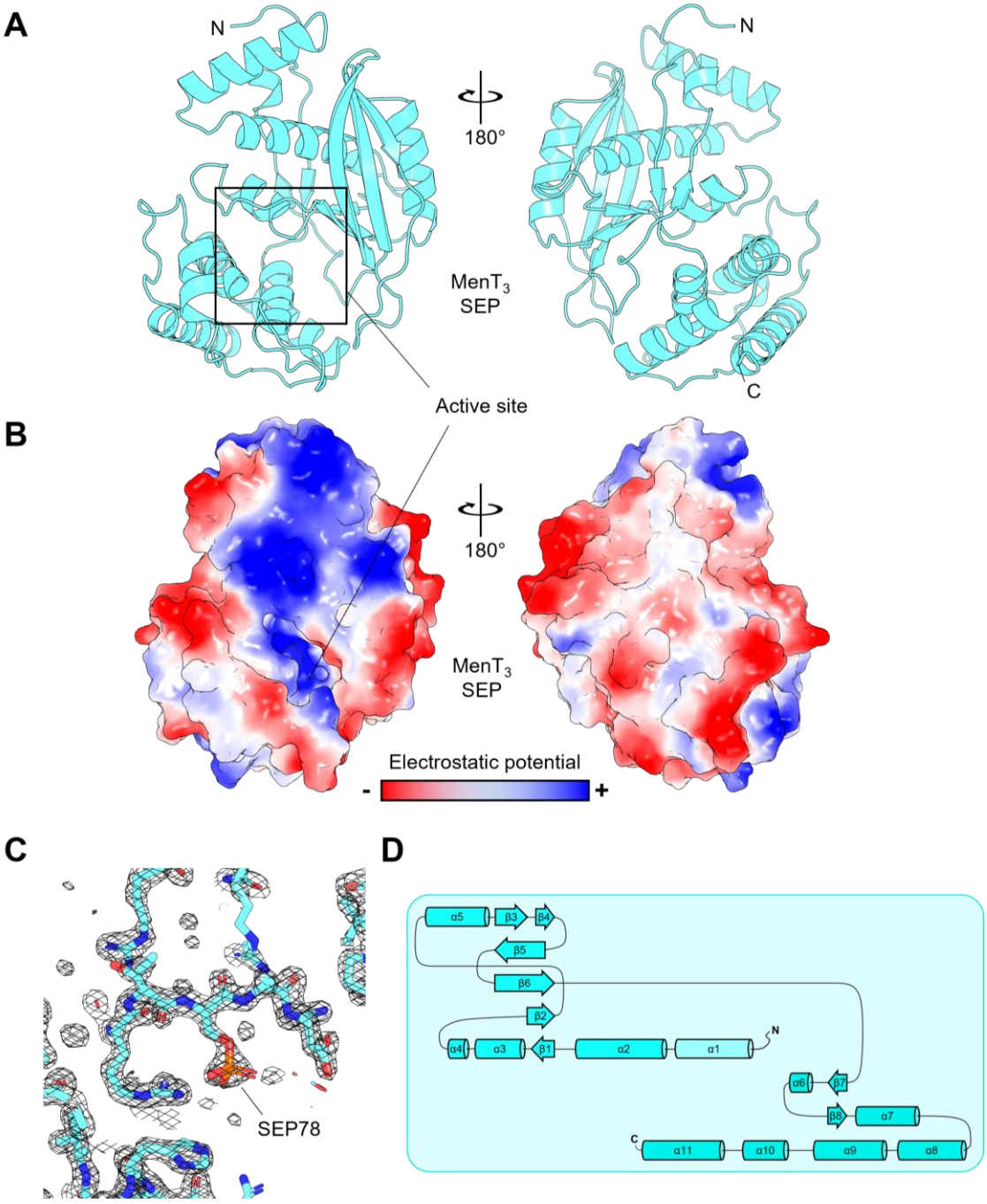


Figure 5.2. Crystal structure of the MenT$_3$ SEP toxin. (A) Structure of monomeric MenT$_3$ SEP with front and back views rotated on *y* by 180°, displayed as aquamarine cartoon representations. **(B)** MenT$_3$ SEP surface electrostatics viewed as in (A); red represents electronegative potential, blue represents electropositive potential. **(C)** Close-up view of the MenT$_3$ SEP phosphoserine shown with a 2Fo-Fc electron density map contoured to 2σ. **(D)** Topology of the MenT$_3$ SEP protein.

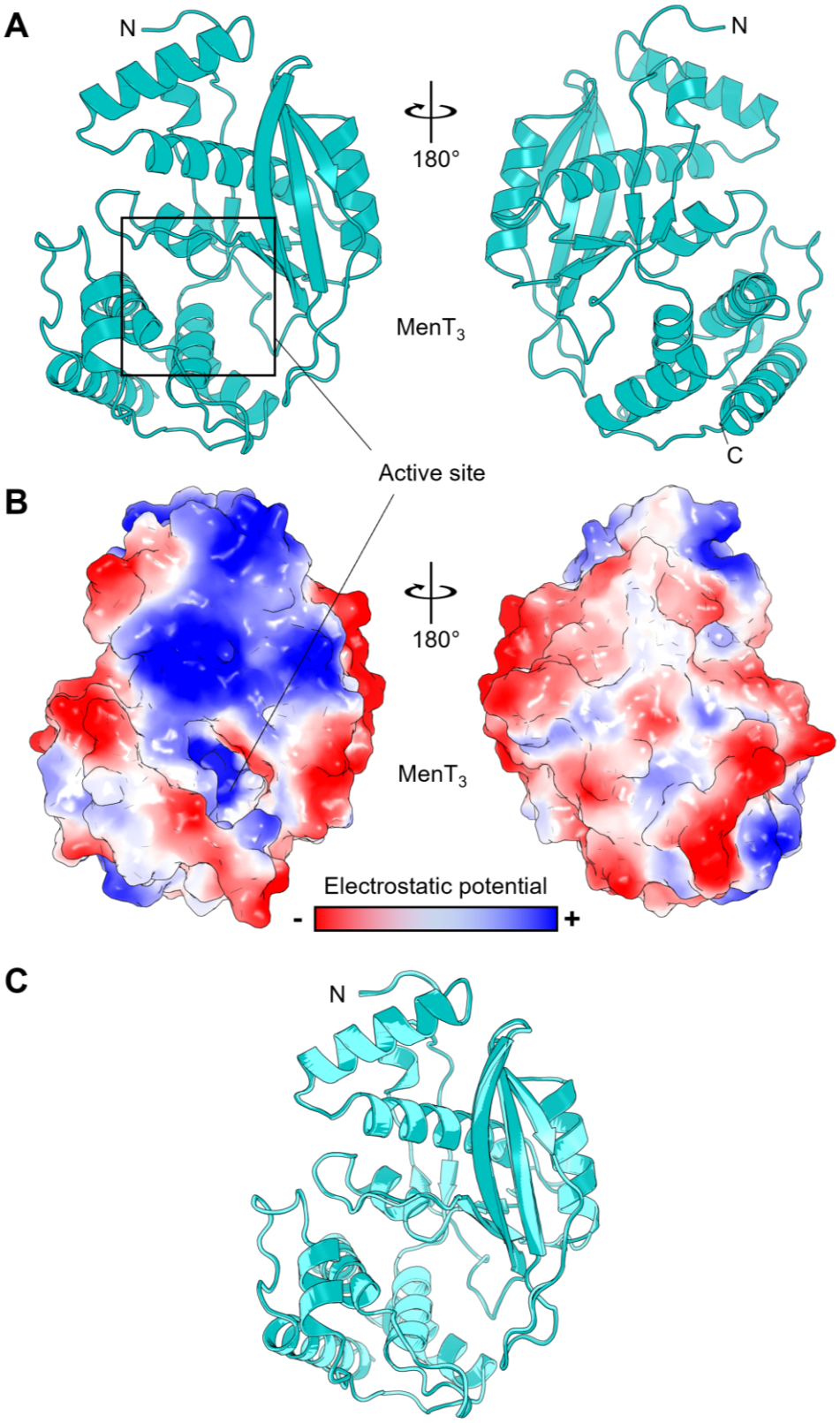


Figure 5.3. Crystal structure of the MenT$_3$ toxin. (A) Structure of monomeric MenT$_3$ with front and back views rotated on *y* by 180°, displayed as teal cartoon representations. **(B)** MenT$_3$ surface electrostatics viewed as in (A); red represents electronegative potential, blue represents electropositive potential. **(C)** Alignment of the MenT$_3$ and MenT$_3$ SEP crystal structures, displayed as teal and aquamarine cartoon representations, respectively.

5.3.1. Conserved active site residues are essential for $MenT_3$ toxicity

The anomalous presence of the phosphoserine at S78 of the $MenT_3$ SEP crystal structure was intriguing, yet the cause of this modification was unknown. The cognate $MenA_3$ antitoxin was proposed as potentially responsible, possibly functioning as a kinase. This rationale stemmed from the different expression strains used to produce protein for crystallography: the $MenT_3$ SEP crystal structure originated from an *E. coli* ER2566 pRARE pPF656 pTRB517 strain co-expressing $MenA_3$ and $MenT_3$; the non-phosphorylated $MenT_3$ structure was derived from ER2566 pRARE pTRB517, where $MenT_3$ was expressed alone (See, *5.3*). It was hypothesised that S78 phosphorylation might be mediated via $MenA_3$ co-expression, potentially contributing to antitoxicity by causing steric hindrance within the putative active site or changing the electrostatic charge to disrupt $MenT_3$-NTP substrate binding. Consequently, the S78 phosphoserine, and therefore possibly an S78A substitution mutant, might detrimentally impact $MenT_3$ toxicity. Accordingly, the S78 of $MenT_3$ was engineered for alanine substitution by SDM (Materials and Methods, *2.3.2.5*). Initial attempts to generate $MenT_3^{(S78A)}$ by SDM using pPF657 (*$menT_3$* WT) as a template repeatedly failed. As a result, pTRB517 was instead used to first generate pTRB630 by SDM, encoding His_6-SUMO-tagged $MenT_3^{(S78A)}$. This was in turn used as a template to amplify a DNA fragment by PCR encoding untagged *$menT_3^{(S78A)}$*, which was then LIC cloned into pTRB551 to produce pTRB631 for toxicity testing.

ConSurf was also used to calculate $MenT_3$ residue conservation[359]. This output was mapped onto the $MenT_3$ structure, which showed that the distribution of conserved residues was mainly clustered in the electropositive putative active site (Figure 5.4A). In addition, a previous AA sequence alignment of DUF1814 proteins had highlighted conserved residues which were shown to be essential to AbiEii toxicity in *S. agalactiae*[83]. ConSurf plots confirmed that several of these residues were conserved and located within the predicted $MenT_3$ active site (Figure 5.4B). The presence of structurally conserved AAs is a good indicator of functional relevance. Therefore, to establish their relative importance to $MenT_3$ toxicity, residues K61, D80, Q185, Q188, K189, and D211 were also engineered for alanine substitution by SDM. Plasmid pPF657 was used as a template to generate pTRB559 (*$menT_3^{(K61A)}$*), pTRB591 (*$menT_3^{(D80A)}$*), pTRB560 (*$menT_3^{(Q185A)}$*), pTRB561 (*$menT_3^{(Q188A)}$*), pTRB562 (*$menT_3^{(K189A)}$*), and pTRB592 (*$menT_3^{(D211A)}$*).

E. coli DH5α were separately co-transformed with each of the $MenT_3$ mutant plasmids and the pTA100 empty vector and tested by toxicity assays, alongside DH5α pTA100 pPF657 as a positive control of toxicity (Figure 5.5A). Of the samples tested, both $MenT_3^{(Q185A)}$ and

$MenT_3^{(Q188A)}$ retained toxicity comparable to $MenT_3$ WT. Interestingly, growth was also inhibited in *E. coli* expressing $MenT_3^{(S78A)}$, suggesting that residues S78, Q185 and Q188 are not functionally relevant to the $MenT_3$ mode of action. In contrast, toxicity was abolished for $MenT_3^{(K61A)}$, $MenT_3^{(D80A)}$, $MenT_3^{(K189A)}$, and $MenT_3^{(D211A)}$, highlighting these residues as essential for $MenT_3$ activity, and further supporting the proposed location for the $MenT_3$ active site (Figure 5.5A).

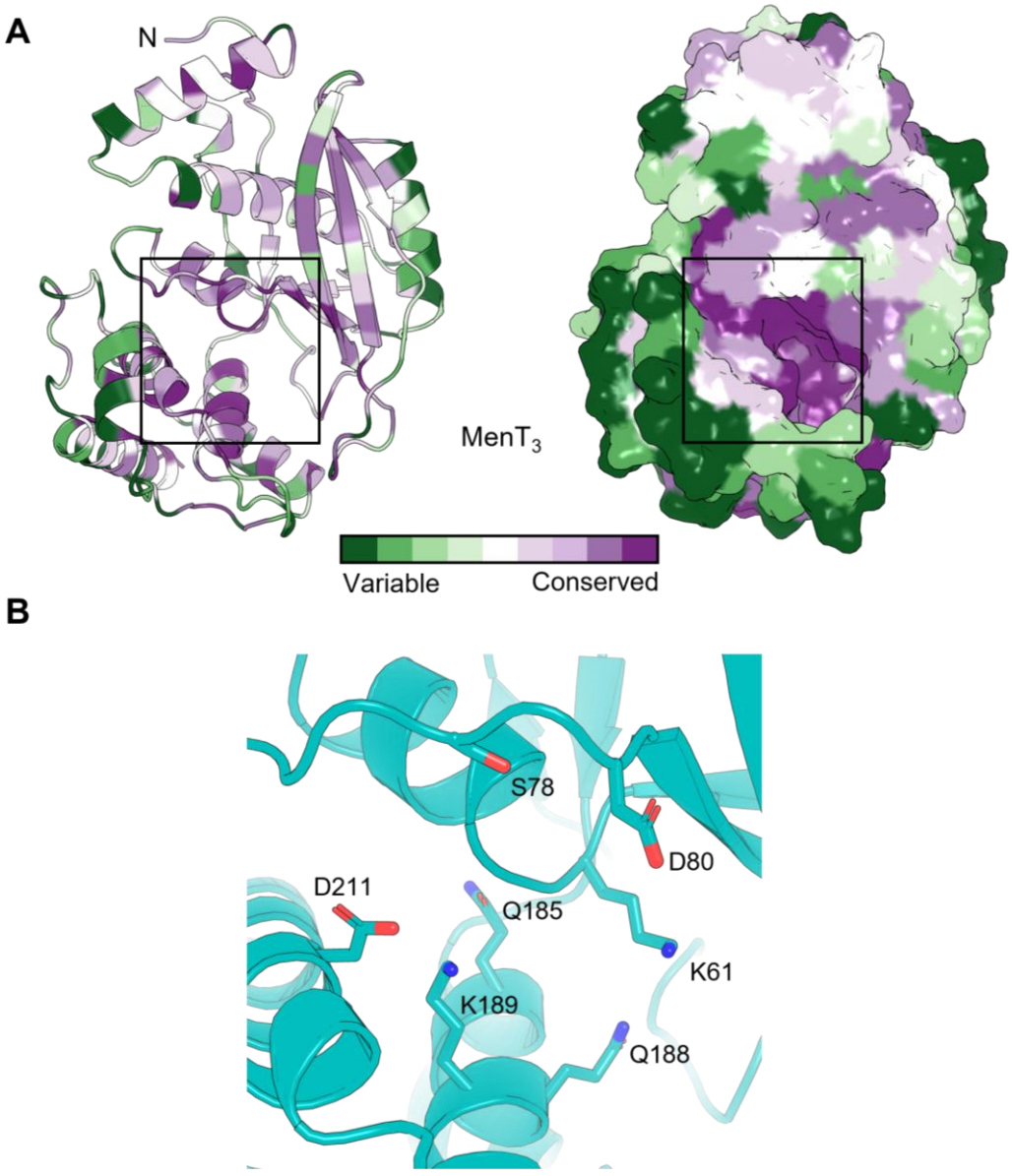


Figure 5.4. Conserved $MenT_3$ residues indicate a putative active site. (A) Conservation plots on the $MenT_3$ toxin structure, displayed as a cartoon representation (left) and surface representation (right), coloured green to purple as per scale. **(B)** Close-up view of the putative toxin active site, as indicated by the boxed regions in (A). $MenT_3$ residues K61, S78 (subject to phosphorylation in $MenT_3$ SEP), D80, Q185, Q188, K189 and D211 are highlighted as stick representations.

5.3.2. $MenT_3$ S78 is essential to $MenA_3$ antitoxicity

In order to round off functionally characterising the identified conserved $MenT_3$ residues, the various $MenT_3$ mutant constructs were next tested in *E. coli* via antitoxicity assays. The MenA-MenT systems were originally predicted to be non-interacting type IV TA modules[11], therefore specific $MenT_3$ residues should theoretically be irrelevant to antitoxicity. However, the canonical class for this family of TA systems had not yet been determined. As such, it could not be ruled out that they may act via a type II mechanism, where the MenA antitoxin would directly interact with the cognate MenT toxin to inhibit toxicity. The implied functional relevance of conserved $MenT_3$ residues might translate to a biologically important site for direct or indirect $MenA_3$ interactions, or possibly protein-protein binding by an unknown antitoxicity cofactor. Indeed, the phosphorylation of $MenT_3$ S78, putatively linked to $MenA_3$ co-expression, supported this prediction.

Accordingly, *E. coli* DH5α were separately co-transformed with pPF656 (*$menA_3$*) and each of the $MenT_3$ mutant plasmids, then assessed by antitoxicity assays (Figure 5.5B). The DH5α pPF656 pPF657 strain was also tested as a positive control of antitoxicity. For almost all of the mutants assayed, the presence of $MenA_3$ was sufficient to neutralise any previously demonstrated toxicity and rescue growth, indicating no functional relevance to $MenA_3$ antitoxicity (Figure 5.5B). The exception was $MenT_3^{(S78A)}$, which continued to inhibit growth despite $MenA_3$ co-expression. This was particularly notable given the potent antitoxicity demonstrated by $MenA_3$ against $MenT_3$ WT, where even uninduced $MenA_3$ was sufficient to neutralise toxicity (Figure 5.5B). These observations indicated that $MenT_3$ S78 is a specific and essential residue for the antitoxic activity of $MenA_3$; the S78 alanine substitution likely prevented $MenA_3$-mediated S78 phosphorylation, and consequently prevented $MenT_3$ toxin neutralisation.

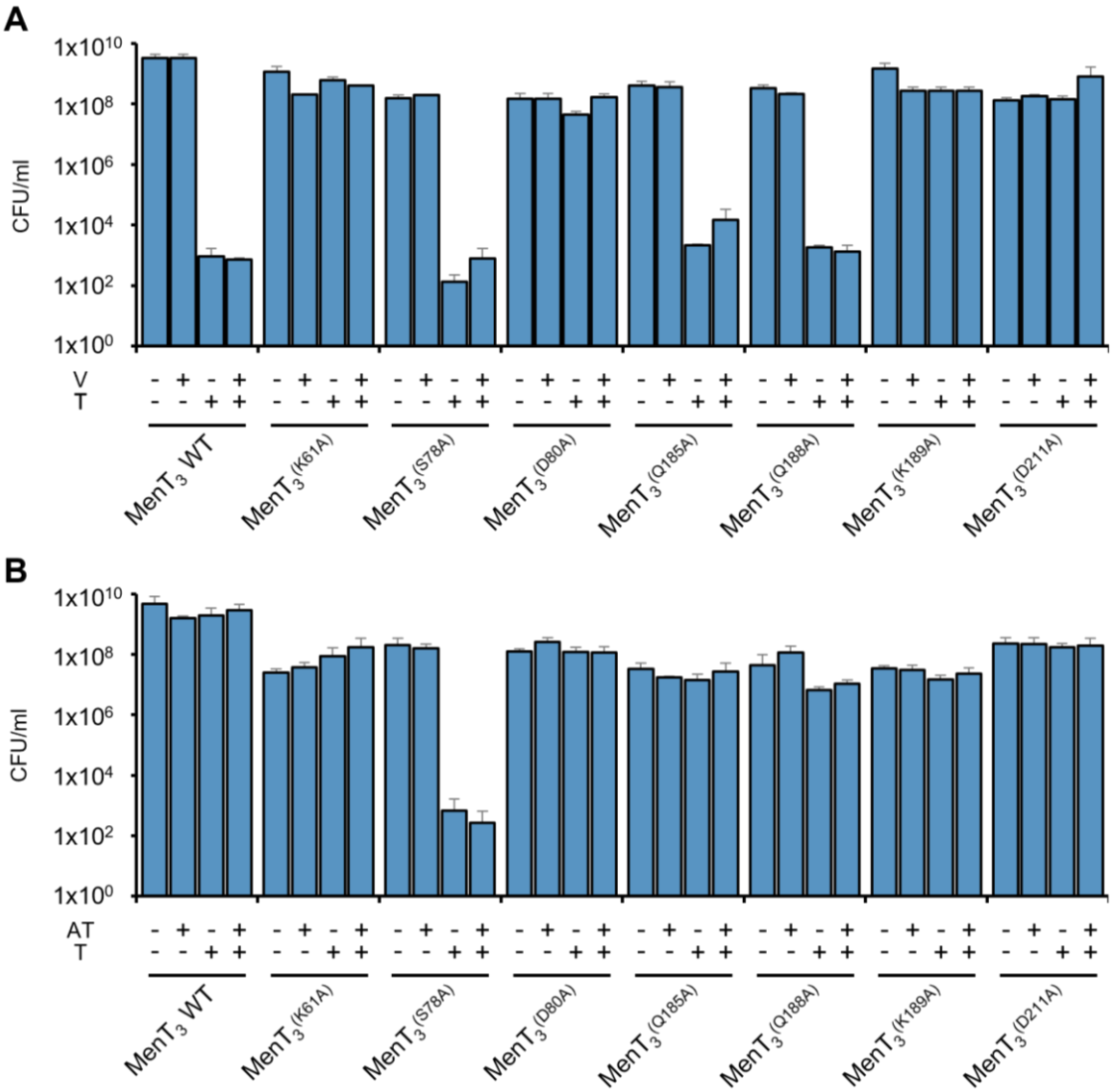


Figure 5.5. Substitutions of conserved $MenT_3$ residues within the putative active site abolish toxicity. (A) Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with the pTA100 empty vector and either $MenT_3$ WT (pPF657), K61A (pTRB559), S78A (pTRB631), D80A (pTRB591), Q185A (pTRB560), Q188A (pTRB561), K189A (pTRB562) or D211A (pTRB592) substitution constructs. Overnight cultures were re-seeded into fresh LB supplemented with Ap, Sp and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Ap and Sp, and with or without D-glu, L-ara and IPTG for repression of toxin expression, induction of toxin expression, and induction of antitoxin expression, respectively. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "V" = vector, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates). **(B)** Endpoint viable count antitoxicity assays as in (A), with *E. coli* DH5α transformed with pPF656 (*menA*$_3$) and either the $MenT_3$ WT or $MenT_3$ substitution constructs. "AT" = antitoxin, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

5.4. $MenT_4$ crystal structure

To obtain the $MenT_4$ crystal structure, *E. coli* ER2566 was first transformed with pRARE and pTRB544. $MenT_4$ was expressed and purified as per Materials and Methods, *2.6* and *2.7*. The expression culture grew at a steady rate, and protein purification proceeded smoothly. However, the overall protein yield was notably poor in comparison to $MenT_3$ expression, even where $MenT_3$ was expressed in the absence of the cognate antitoxin. This was observed initially via a low chromatogram peak after AEC, reflecting relatively low levels of protein, and confirmed by an even lower peak after SEC and by SDS-PAGE analysis. Initial attempts to concentrate the sample to 12 mg/ml caused protein to precipitate out of solution, with the highest achievable concentration prior to protein precipitation being 6 mg/ml. Fortunately, this proved sufficient to set a limited number of crystal screens, and remarkably, the crystals only formed in the identical condition used for $MenT_3$ crystallisation (Materials and Methods, *2.8*). The successful condition was optimised for $MenT_4$ crystals, which were subsequently used to solve the $MenT_4$ structure by *ab initio* methods to 1.23 Å (Materials and Methods, *2.9*) (Table 5.2 and Figure 5.6).

$MenT_4$ features an overall similar architecture to $MenT_3$; it is a bi-lobed globular protein with a large central cavity separating the NTD and CTD (Figure 5.6A). Surface electrostatics show distinct patches of electropositive potential, comparable to but not as extensive as $MenT_3$, which lead to a similarly positioned region predicted to be the toxin active site (Figure 5.6B). A brief analysis of density at the conserved $MenT_4$ S67 showed no evidence for a phosphoserine comparable to $MenT_3$ SEP. This was consistent with the expression method used to produce $MenT_4$ protein for crystallisation, where only $MenT_4$ was expressed, without $MenA_4$. PDBsum was used to generate a $MenT_4$ protein topology diagram[387], which showed broadly similar secondary structure motifs to $MenT_3$, characterised by a helical/anti-parallel β-sheet NTD and a predominantly helical CTD (Figure 5.6C).

Table 5.2. Crystallographic data collection and refinement statistics

	$MenT_4$ Native	$MenT_1$ Native	$MenA_1$:$MenT_1$ Native
Data collection			
PDB ID Code	6Y56	-	-
Beamline	Diamond I24	Diamond I24	Diamond I04
Wavelength, Å	0.9781	0.9795	0.9795
Resolution range, Å	42.23 – 1.23 (1.27 – 1.23)	43.84 – 1.65 (1.71 – 1.65)	43.68 – 1.44 (1.49 – 1.44)
Space group	$P2_1$	I422	$P12_11$
Unit cell, *a b c* (Å), α β γ (°)	42.3 57.8 54.7, 90.0 92.3 90.0	123.9 123.9 118.3 90.0 90.0 90.0	45.9 81.7 64.8 90 107.8 90
Total reflections	149653	5659100	1110358
Unique reflections	75996 (7206)	55309 (5434)	82208 (8166)
Multiplicity	2.0	102.3	13.5
Completeness (%)	98.80 (88.97)	90.68 (14.59)	99.8 (98.70)
Mean I/ σ (I)	7.0	8.95	13.07
R_{merge}	0.060 (0.926)	1.818 (-20.7)	0.088 (2.152)
R_{meas}	0.085 (1.310)	1.827 (-20.8)	0.091 (2.240)
$CC_{1/2}$	0.996 (0.294)	0.994 (0.292)	0.999 (0.632)
Refinement			
R_{work}	0.1840 (0.3174)	0.2247 (0.7550)	0.1706 (0.3348)
R_{free}	0.1950 (0.3352)	0.2333 (0.9489)	0.1899 (0.3398)
No. of non-hydrogen atoms	2649	3077	3679
macromolecules	2322	2957	3335
molvent	327	120	344
Protein residues	292	384	429
RMSD (bonds, Å)	0.005	0.013	0.010
RMSD (angles, °)	0.830	1.41	1.12
Ramachandran favored (%)	98.28	97.88	98.56
Ramachandran allowed (%)	1.72	2.12	1.44
Ramachandran outliers (%)	0.00	0.00	0.00
Average B-factor	20.6	43.9	37.7
macromolecules	19.3	44.0	37.4
solvent	29.9	41.5	40.4

Statistics for the highest-resolution shell are shown in parentheses

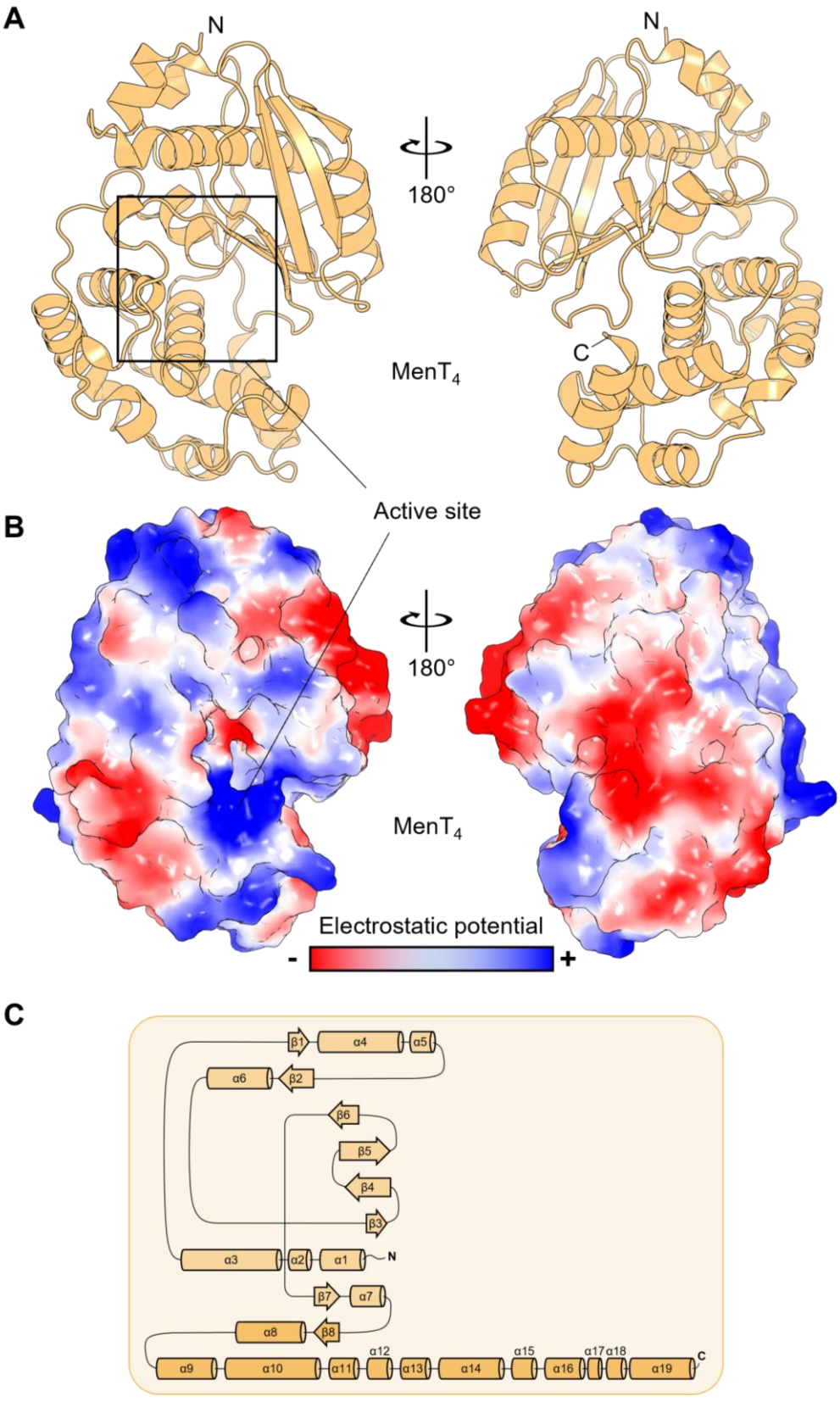


Figure 5.6. Crystal structure of the MenT$_4$ toxin. (A) Structure of monomeric MenT$_4$ with front and back views rotated on *y* by 180°, displayed as light orange cartoon representations. **(B)** MenT$_4$ surface electrostatics viewed as in (A); red represents electronegative potential, blue represents electropositive potential. **(C)** Topology of the MenT$_4$ protein.

5.4.1. Superposition of $MenT_3$ and $MenT_4$ highlights a conserved active site

A ConSurf analysis was again performed, this time to calculate $MenT_4$ residue conservation[359]. Similar to $MenT_3$, conserved residues predominantly clustered in the electropositive central cavity, suggesting a conserved $MenT_3$ and $MenT_4$ active site (Figure 5.7A). Aligning $MenT_3$ and $MenT_4$ by sequence gave a poor RMSD of 13.420 Å, between 1304 atoms, consistent with the relatively low AA sequence identity between the two proteins (26.4% (EMBOSS Stretcher)). However, performing a sequence-independent superposition improved the RMSD to 4.661 Å, between 1069 atoms (Figure 5.7B). This shows that $MenT_3$ and $MenT_4$ share the same overall fold and secondary structural motifs, but relative numbers, positions and lengths of secondary structural elements can vary. $MenT_3$ and $MenT_4$ were next aligned by AA sequence using Jalview[388], via the Clustal W plugin[389], which highlighted $MenT_4$ S67, D69, K171 and D186 as conserved residues corresponding to functionally relevant $MenT_3$ residues S78, D80, K189 and D211, respectively. A close-up of the superposed active site shows that the highlighted $MenT_4$ residues occupy similar positions to those of the homologous $MenT_3$ residues, suggesting that they may also have similar functional relevance (Figure 5.7C).

5.4.2. Conserved $MenT_4$ active site residues are required for toxicity in *E. coli*

The conserved S67, D69, K171 and D186 $MenT_4$ active site residues were carried forward for mutagenesis studies due to the functional relevance of the equivalent $MenT_3$ residues (Figure 5.8). Each AA was separately substituted to an alanine via SDM using pPF659 (*$menT_4$* WT) as a template, generating pTRB618 (*$menT_4^{(S67A)}$*), pTRB619 (*$menT_4^{(D69A)}$*), pTRB620 (*$menT_4^{(K171A)}$*), and pTRB621 (*$menT_4^{(D186A)}$*). These were then used to co-transform *E. coli* DH5α alongside either pTA100 or pPF658 (*$menA_4$*), with the resulting strains functionally tested by toxicity and antitoxicity assays (Figure 5.8).

Consistent with the equivalent $MenT_3$ mutants, toxicity was abolished for $MenT_4^{(D69A)}$, $MenT_4^{(K171A)}$, and $MenT_4^{(D186A)}$, indicating that these residues are essential to the $MenT_4$ mode of action (Figure 5.8A). Interestingly, $MenT_4^{(S67A)}$ (corresponding to $MenT_3$ S78) also saw toxicity abolished (Figure 5.8A). This contrasted with $MenT_3^{(S78A)}$, which retained toxicity comparable to $MenT_3$ WT (Figure 5.5A). These conflicting phenotypes are intriguing, suggesting that this conserved serine has evolved divergently in MenT toxins to perform separate yet vital functions; for $MenT_3$, S78 appears to be a target for $MenA_3$ phosphorylation and antitoxicity, whereas for $MenT_4$, S67 is essential for toxicity. Antitoxicity assays were also

performed to complete the functional characterisation of $MenT_4$ mutants (Figure 5.8B). However, these were rendered redundant given the lack of $MenT_4$ mutant toxicity with which to test antitoxin activity against.

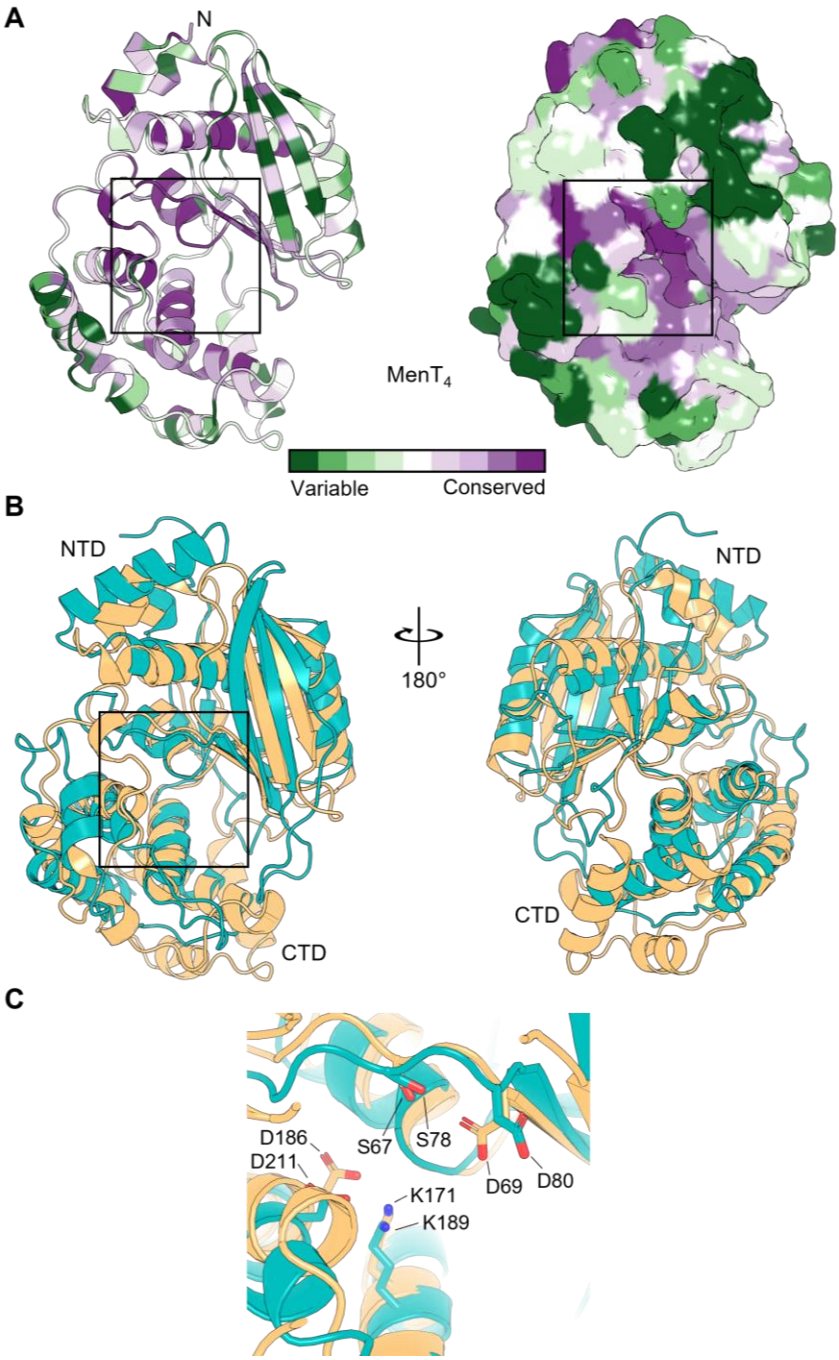


Figure 5.7. $MenT_3$ and $MenT_4$ superposition reveals conserved active site residues. (A) Conservation plots on the $MenT_4$ toxin structure, displayed as a cartoon representation (left) and surface representation (right), coloured green to purple as per scale. **(B)** Superposition of the $MenT_3$ and $MenT_4$ crystal structures with front and back views rotated on *y* by 180°, displayed as teal and light orange cartoon representations, respectively. **(C)** Close-up view of the superposed $MenT_3$ and $MenT_4$ active sites, as indicated by the boxed regions in (A) and (B). Conserved residues are labelled and displayed as stick representations.

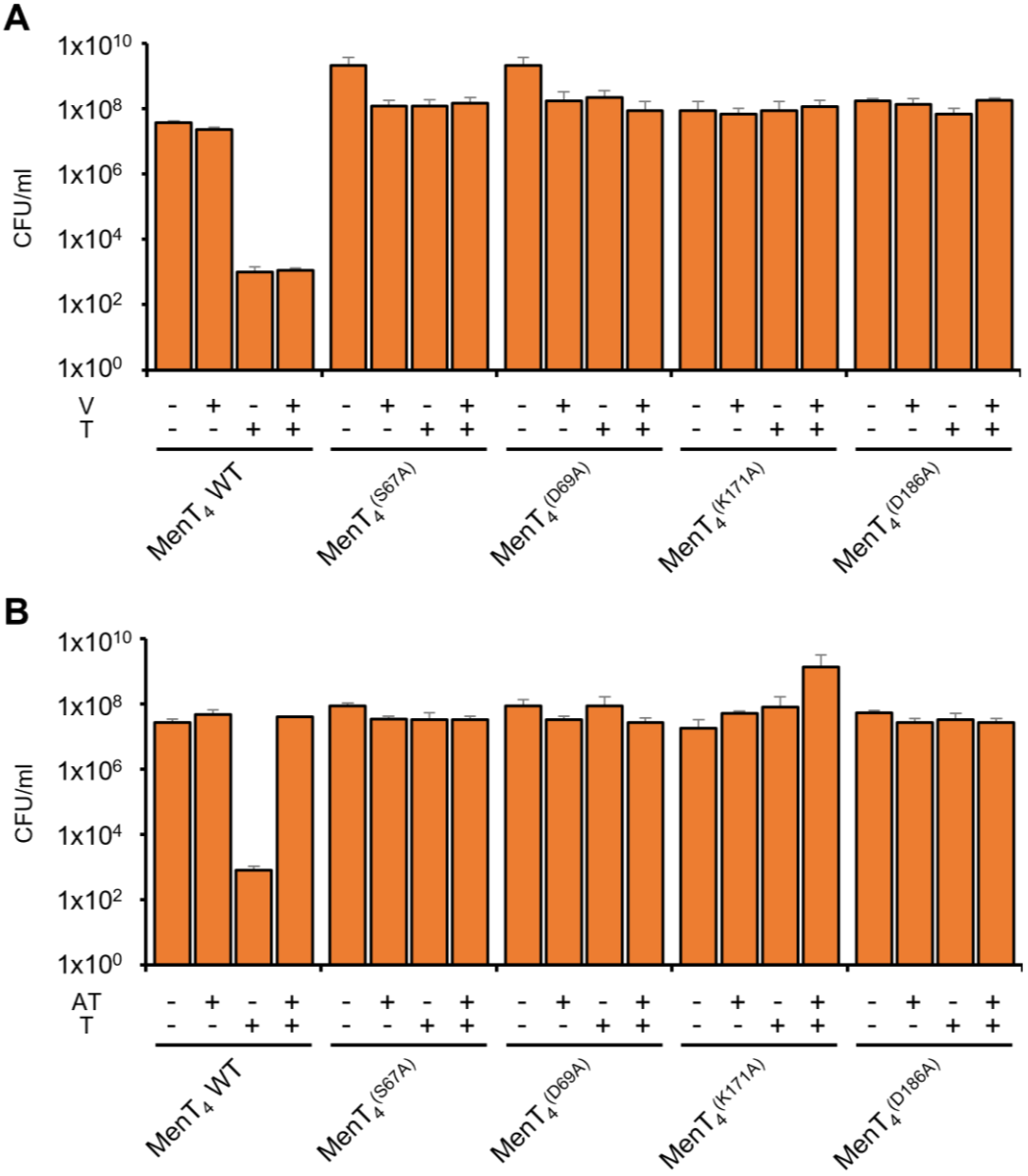


Figure 5.8. Substitutions of conserved MenT$_4$ residues abolish toxicity. (A) Endpoint viable count antitoxicity assays of *E. coli* DH5α transformed with the pTA100 empty vector and either MenT$_4$ WT (pPF659), S67A (pTRB618), D69A (pTRB619), K171A (pTRB620) or D186A (pTRB621) substitution constructs. Overnight cultures were re-seeded into fresh LB supplemented with Ap, Sp and D-glu and grown to mid-log phase. Samples were serially diluted and spotted on M9A plates containing Ap and Sp, and with or without D-glu, L-ara and IPTG for repression of toxin expression, induction of toxin expression, and induction of antitoxin expression, respectively. Plates were incubated at 37 °C for two days, after which they were imaged and colonies counted to determine CFU/ml. "V" = vector, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates). **(B)** Endpoint viable count antitoxicity assays as in (A), with *E. coli* DH5α transformed with pPF658 (*menA$_4$*) and either the MenT$_4$ WT or MenT$_4$ substitution constructs. "AT" = antitoxin, "T" = toxin; "-" and "+" denote expression state. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

5.5. $MenA_1$-$MenT_1$ structural characterisation

Having obtained the crystal structures of the $MenT_3$ and $MenT_4$ toxins and characterised conserved residues within their putative active sites, efforts focused on structurally characterising the recently identified $MenA_1$-$MenT_1$ TA system. The *menA₁* gene was amplified by PCR from *M*. *tuberculosis* H37Rv genomic DNA, and *menT₁* from pET-$MenT_{1\text{-His}}$ (kindly provided by the Genevaux lab), and both were cloned by LIC into pTRB550 to create pTRB617 (*menA₁*) and pTRB629 (*menT₁*). Due to the relatively small size of the $MenA_1$ and $MenT_1$ proteins, at 7.41 kDa and 21.48 kDa, respectively, pTRB617 and pTRB629 were used to separately transform *E. coli* BL21(λDE3)Δ*slyD*, also provided by the Genevaux lab. This strain lacks the *slyD* gene encoding the 20.85 kDa SlyD protein, a common contaminant in purifications from other *E. coli* expression strains. Whilst the two-step affinity chromatography purification methodology used in this study had proved more than sufficient to purify $MenT_3$ and $MenT_4$ proteins, the *E. coli* BL21(λDE3)Δ*slyD* strain was employed as insurance to ensure purity of the final sample.

To begin with, *E. coli* BL21(λDE3)Δ*slyD* pTRB617 and BL21(λDE3)Δ*slyD* pTRB629 were expressed as described previously (Materials and Methods, *2.6*). The potential growth-inhibitory effects of $MenT_1$ expression were not a concern, as the Genevaux lab had reported that $MenT_1$ was non-toxic in *E. coli* (personal communication). Purifications proceeded as per Materials and Methods, *2.7*. However, after the second affinity chromatography step, and prior to SEC purification, the $MenA_1$ antitoxin and $MenT_1$ toxin samples were each divided into two. One sample each of $MenA_1$ and $MenT_1$ were directly mixed and co-incubated overnight at -4 °C. The following morning, this sample ($MenA_1$:$MenT_1$) was concentrated and purified by SEC. Analysis of the corresponding SEC chromatogram peak by SDS-PAGE revealed the presence of purified $MenA_1$ and $MenT_1$, suggesting that these proteins had complexed *in vitro*. The remaining separate $MenA_1$ and $MenT_1$ protein samples were also concentrated and purified by SEC. The resulting SEC chromatogram showed a high yield of purified $MenT_1$ toxin protein, which was confirmed by SDS-PAGE. In contrast, no chromatogram peak was visible following $MenA_1$ SEC, with SDS-PAGE analysis confirming the loss of $MenA_1$ protein. This was puzzling, as AEC chromatograms and SDS-PAGE analysis earlier in the purification process had indicated a high yield of $MenA_1$. Presumably the $MenA_1$ protein was either too unstable, or too small for the resolution capability of the SEC column used. As a result, this sample was discarded, whilst the purified $MenT_1$ and $MenA_1$:$MenT_1$ proteins were carried forward for crystallographic studies.

Crystal screens were performed with both $MenT_1$ and $MenA_1$:$MenT_1$ protein samples concentrated to 12 mg/ml in Crystal buffer (Materials and Methods, *2.8*). The resulting $MenT_1$ and $MenA_1$:$MenT_1$ crystals developed in separate conditions; these were harvested directly from the screens, and X-ray crystallographic data were collected as per Materials and Methods, *2.9*. Similar to $MenT_4$, the $MenA_1$:$MenT_1$ X-ray diffraction dataset proved to be of a sufficiently high resolution to solve the crystal structure *ab initio* using Arcimboldo[351], to a resolution of 1.44 Å (Table 5.2). This structure was in turn used as a search model to solve the crystal structure of $MenT_1$ by MR, to 1.65 Å resolution (Table 5.2) (Materials and Methods, *2.9*).

5.5.1. $MenT_1$ crystal structure

The $MenT_1$ crystal structure showed two $MenT_1$ protomers in the crystallographic unit ($MenT_1$[A] and $MenT_1$[B]), that appeared to bind asymmetrically to each other (Figure 5.9A). As with $MenT_3$ and $MenT_4$, the $MenT_1$[A] protomer is bi-lobed and globular, with the distinct NTD and CTD separated by an electropositive central cavity (Figure 5.9B and C). As expected, $MenT_1$[A] and $MenT_1$[B] aligned well, giving an RMSD of 0.327 Å, between 2166 atoms (Figure 5.9D). The $MenT_1$[A] protein topology was derived from PDBsum[387], which showed less extensive but broadly comparable secondary structure motifs to $MenT_3$ and $MenT_4$ (Figure 5.9E).

5.5.2. $MenT_1$, $MenT_3$, and $MenT_4$ superpositions confirm a MenT family toxin fold

ConSurf was used to calculate the residue conservation of $MenT_1$[A][359], which showed that conserved residues were primarily clustered in the toxin central cavity (Figure 5.10A). $MenT_1$[A] was next aligned to $MenT_3$ and $MenT_4$, which gave a poor RMSD of 13.212 Å, between 1674 atoms, and 13.684 Å, between 585 atoms, respectively. However, sequence-independent superpositions of $MenT_1$[A] with $MenT_3$ and $MenT_4$ gave improved RMSD values of 4.561 Å, between 1510 atoms, and 4.232 Å, between 889 atoms, respectively (Figure 5.10B). Superposing the three $MenT_1$[A], $MenT_3$ and $MenT_4$ structures highlighted that whilst the secondary structure elements exhibit a high degree of variability, most notable in the much smaller $MenT_1$ protein, all three MenT toxins share a conserved toxin fold (Figure 5.10C). A close-up view of the superposed active sites revealed the presence of $MenT_1$ D41, K137 and D152, which overlayed well with the corresponding conserved $MenT_3$ and $MenT_4$ active site residues (Figure 5.10C). Furthermore, whilst no serine is present in $MenT_1$

comparable to $MenT_3$ S78 or $MenT_4$ S67, $MenT_1$ T39 was instead identified which took up a similar position recessed towards the back of the putative $MenT_1$ active site (Figure 5.10C). Given that $MenT_3$ S78 appeared to be a potential phosphorylation target of $MenA_3$, $MenT_1$ T39 might similarly be a serine/threonine (Ser/Thr) kinase target.

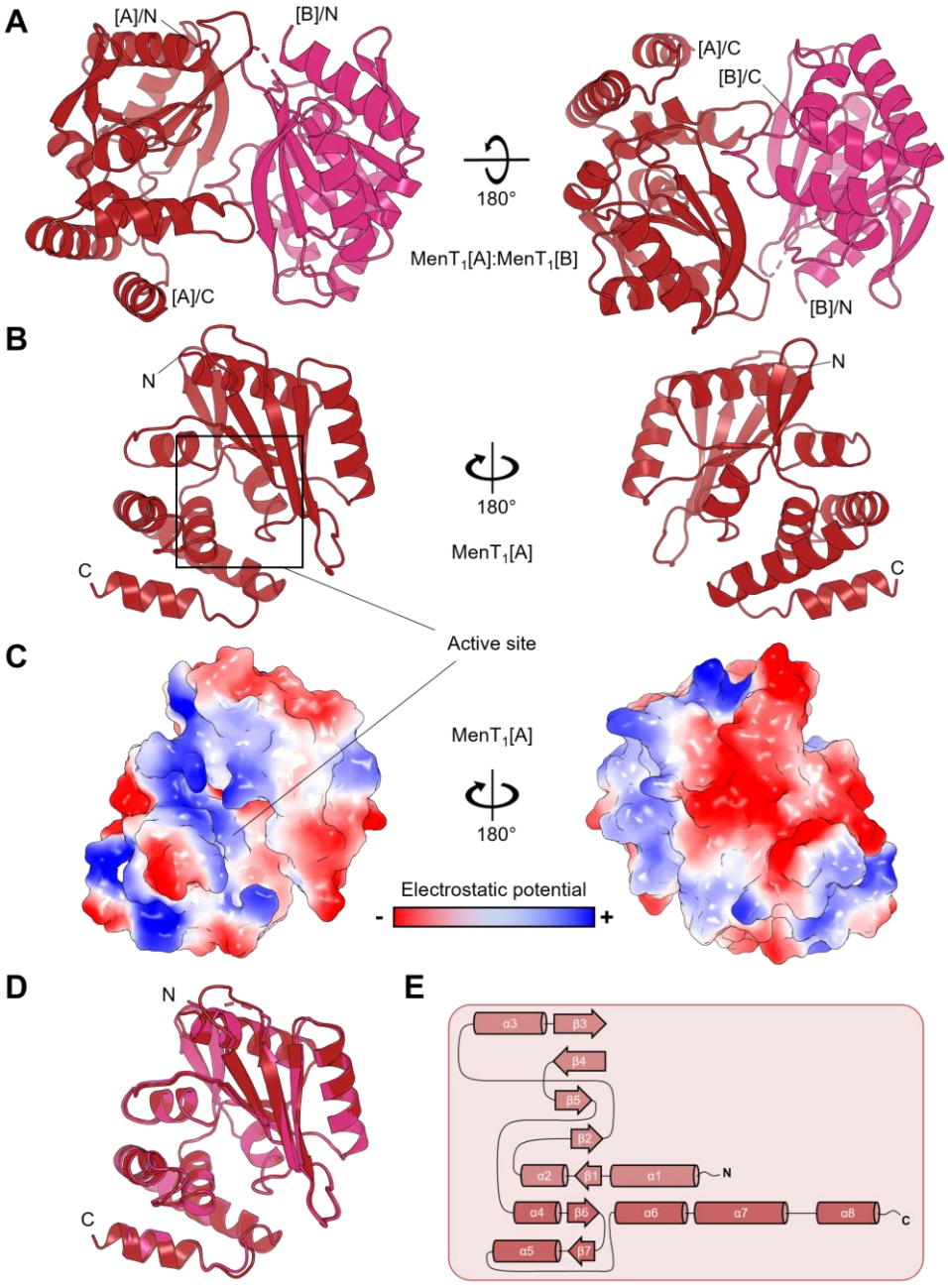


Figure 5.9. Crystal structure of the $MenT_1$ toxin. $MenT_1$ protomer A ([A]) is coloured fire brick, $MenT_1$ protomer B ([B]) is coloured warm pink; the structure is displayed as a cartoon representation. **(A)** Two views of the $MenT_1$ crystal structure, rotated on *x* by 180°. **(B)** Front and back views of the $MenT_1$[A] protomer, rotated on *y* by 180°. **(C)** $MenT_1$[A] surface electrostatics viewed as in (B); red represents electronegative potential, blue represents electropositive potential. (**D**) Alignment of $MenT_1$[A] and $MenT_1$[B]. **(E)** Topology of the $MenT_1$[A] protomer.

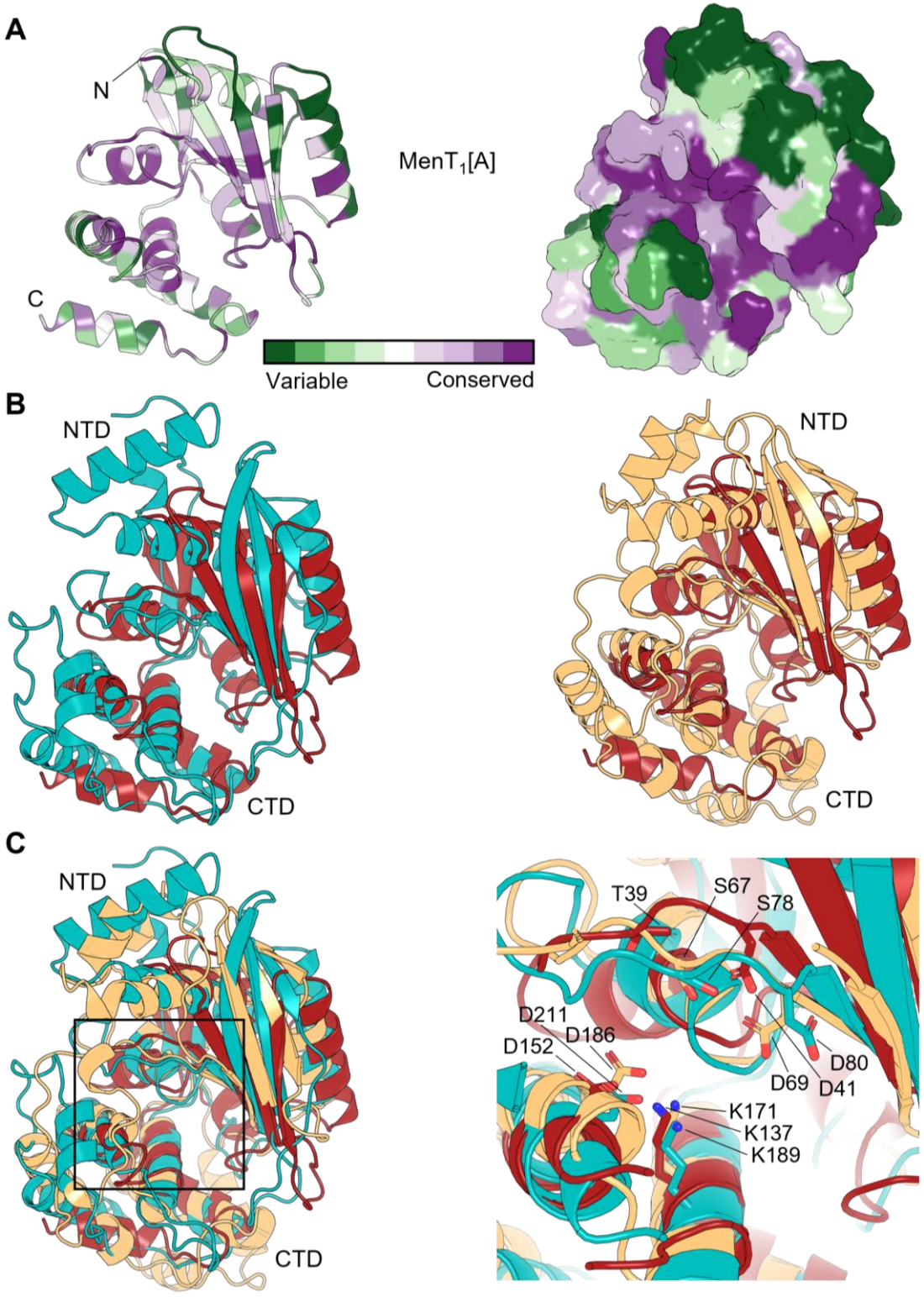


Figure 5.10. MenT$_1$ shares a conserved active site with MenT$_3$ and MenT$_4$. (A) Conservation plots on the MenT$_1$[A] toxin structure, displayed as a cartoon representation (left) and surface representation (right), coloured green to purple as per scale. **(B)** Superposition of MenT$_1$[A] with MenT$_3$ (left) and MenT$_4$ (right), displayed as fire brick, teal, and light orange cartoon representations, respectively. **(C)** Left: superposition of MenT$_1$[A], MenT$_3$ and MenT$_4$. Right: close-up view of the superposed active sites, as indicated by the boxed region in (C, left). Conserved residues are labelled and displayed as stick representations.

5.5.3. $MenA_1$:$MenT_1$ crystal structure

The $MenA_1$:$MenT_1$ crystal structure features two $MenT_1$ protomers ($MenT_1$[A] and $MenT_1$[B]) bound to a $MenA_1$ antitoxin monomer to form a $MenT_1$:$MenA_1$:$MenT_1$ complex (Figure 5.11A). One gap and two gaps are present in the protein chains of $MenT_1$[A] and $MenT_1$[B], respectively, occurring in external flexible loops. The $MenT_1$[A] and $MenT_1$[B] protomers align well, giving an RMSD of 0.494 Å, between 1988 atoms, indicating a near-identical structural architecture (Figure 5.11B, left). Aligning $MenT_1$[A] from the $MenA_1$:$MenT_1$ crystal structure to $MenT_1$[A] from the $MenT_1$ crystal structure gives a similarly good RMSD of 0.462 Å, between 2182 atoms (Figure 5.11B, right).

The $MenA_1$ antitoxin protein consists of a large central helix (α1) linked by a flexible loop to a smaller C-terminal helix (α2). The α2 helix folds back, partially towards the α1 helix centre, yet angled outwards, forming an asymmetrical, twisted "v" shape (Figure 5.11A). Only 51 AAs were resolved in the $MenA_1$ crystal structure. The 17 residues for which there was no electron density comprised the remainder of the $MenA_1$ C-terminus, which is orientated towards the exterior of the crystal structure (Figure 5.11A). The probable positioning of these unresolved residues suggests that they do not interact with either of the $MenT_1$ protomers. These unbound residues therefore likely form an unstable "tail" jutting out of the crystal structure complex, where they are presumably unnecessary for $MenA_1$:$MenT_1$ complex formation. The resolved $MenA_1$ antitoxin binds in opposing directions across the same face of each $MenT_1$ protomer (Figure 5.12A and B). Close-up views of the putative active site of $MenT_1$[A] and $MenT_1$[B] show how formation of the $MenA_1$:$MenT_1$ complex would occlude substrate access, inhibiting toxin activity and thereby neutralising toxicity (Figure 5.13).

As mentioned earlier, $MenA_1$ had previously been identified as a possible toxin by the TASmania database, most closely matching the type I SymE RNase toxin[386]. SymE is itself interesting due to its predicted structural homology to type II MazE antitoxins[57]. This initially seemed like an odd match; the dual-helix crystal structure of $MenA_1$ suggests no RNase capability, yet the proposed similarity to a type I RNase toxin which is itself related to type II DNA-binding antitoxins was intriguing. No SymE structure has so far been solved, therefore Phyre2 and AlphaFold were used to predict and model separate SymE structures[361,362]. A third SymE structure was also assessed, generated by Kawano *et al*. using the SWISS-MODEL homology modelling server[57,390]. All three SymE models exhibited starkly different secondary structure motifs, which raised early concerns about the validity of the predicted structures. Nevertheless, the $MenA_1$ crystal structure was aligned to each SymE model in turn, giving

RMSD scores of 1.226 Å, between 52 atoms (Phyre2), 0.990 Å, between 52 atoms (AlphaFold), and 0.611 Å, between 48 atoms (SWISS-MODEL) (Figure 5.14). The alignments occurred only over very short regions, between a very limited subset of atoms, whilst a closer look showed the secondary structure elements differ greatly between $MenA_1$ and the three SymE structures (Figure 5.14). This indicates that at the structural level there is very little similarity evident between $MenA_1$ and SymE, either implying that the earlier prediction of possible homology was erroneous, or highlighting the vast evolutionary variance that is possible between distantly related proteins. Either way, the alignments suggest that $MenA_1$ does not share the same structural characteristics as SymE, likely reflected in the divergent antitoxic mode of action.

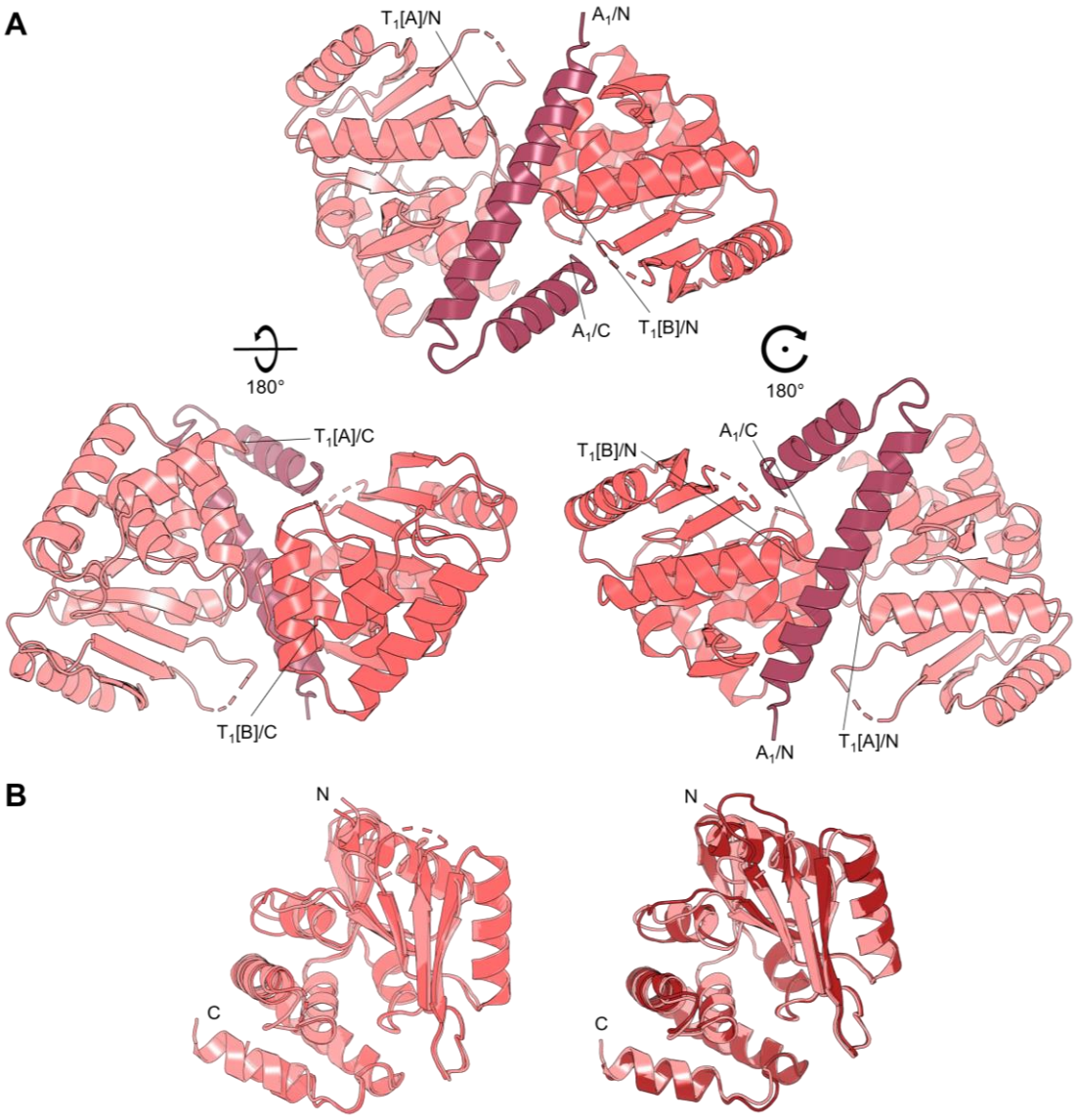


Figure 5.11. Crystal structure of the MenA$_1$:MenT$_1$ TA complex. (A) Different views of the MenA$_1$:MenT$_1$ crystal structure, rotated on *x* by 180° (bottom, left) and on *z* by 180° (bottom, right). MenA$_1$ is coloured raspberry, MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. **(B)** Left: alignment of the MenT$_1$[A] and MenT$_1$[B] protomers from the MenA$_1$:MenT$_1$ crystal structure. Right: alignment of the MenT$_1$[A] protomers from the MenT$_1$ and MenA$_1$:MenT$_1$ crystal structures, coloured fire brick and salmon, respectively.

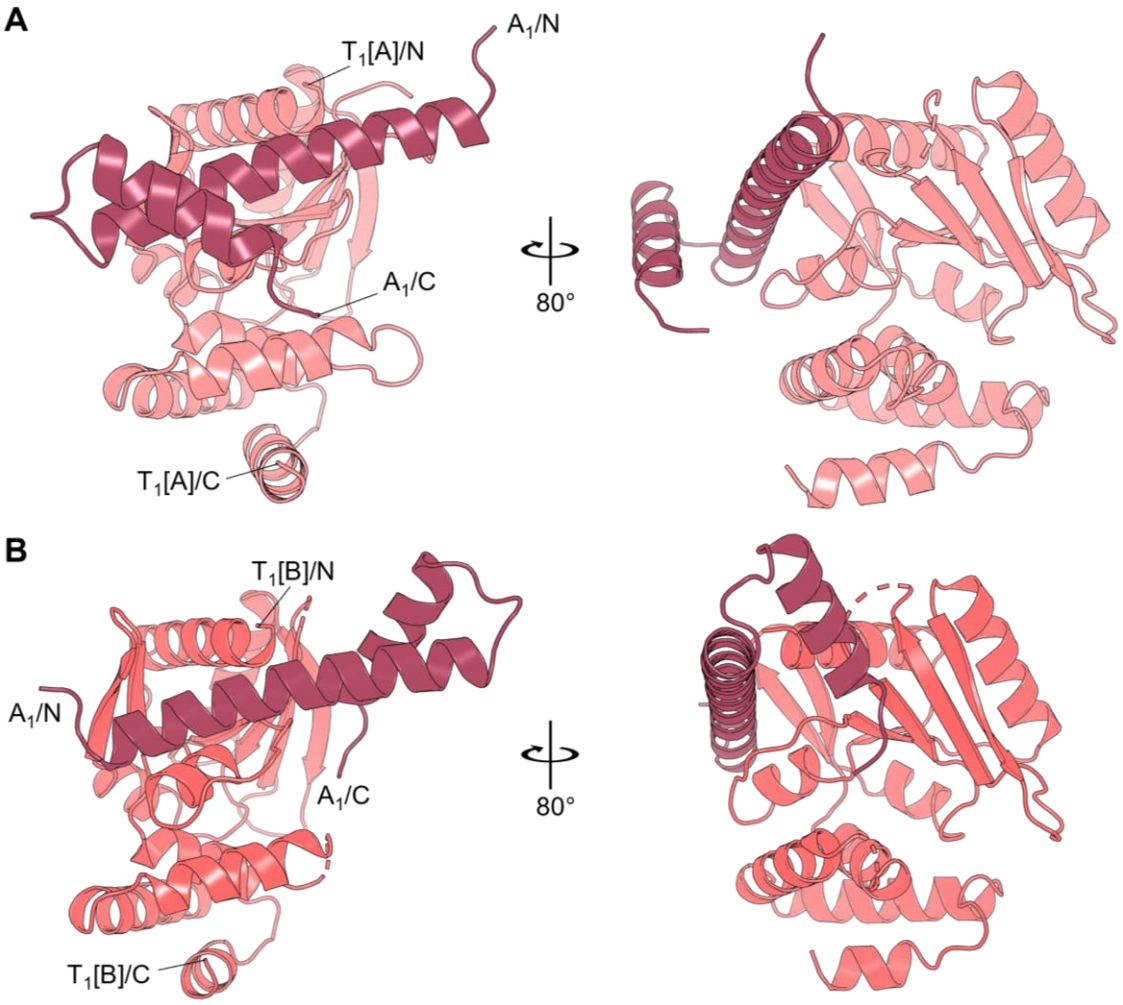


Figure 5.12. MenA$_1$ binds MenT$_1$ protomers asymmetrically. MenA$_1$ is coloured raspberry, MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. **(A)** Views of MenA$_1$:MenT$_1$[A] binding, rotated on *y* by -80°. **(B)** Views of MenA$_1$:MenT$_1$[B] binding, rotated on *y* by -80°.

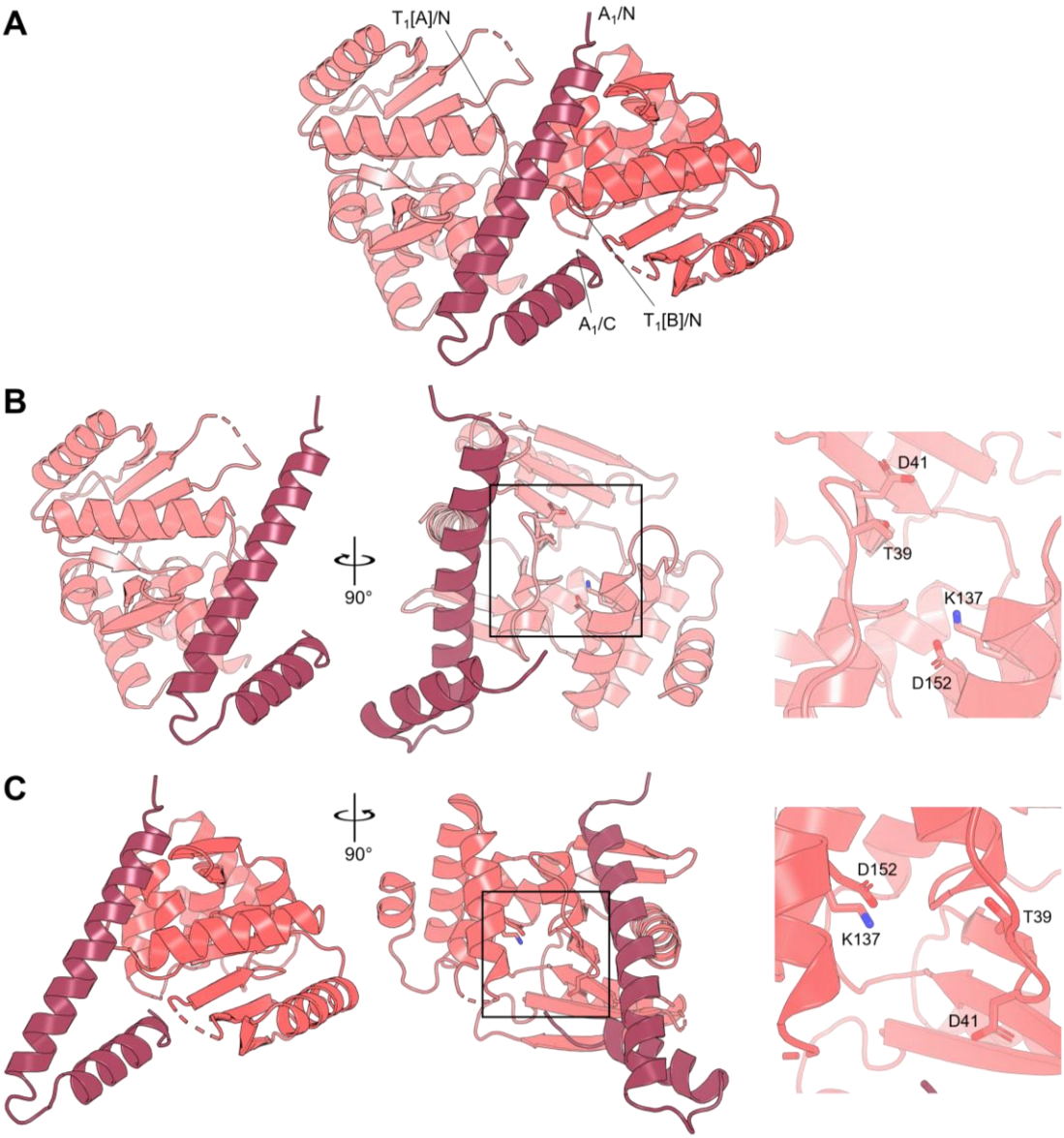


Figure 5.13. MenA$_1$ binding to both MenT$_1$ protomers blocks access to the toxin active site. (A) MenA$_1$ is coloured raspberry, MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. **(B)** View of MenA$_1$:MenT$_1$[A] as in (A), with an alternate view rotated on *y* by -90°. A close-up view of the putative toxin active site is also shown, as indicated by the boxed region in (B, middle). **(C)** View of MenA$_1$:MenT$_1$[B] as in (A), with an alternate view rotated on *y* by 90°. A close-up view of the putative toxin active site is also shown, as indicated by the boxed region in (C, middle).

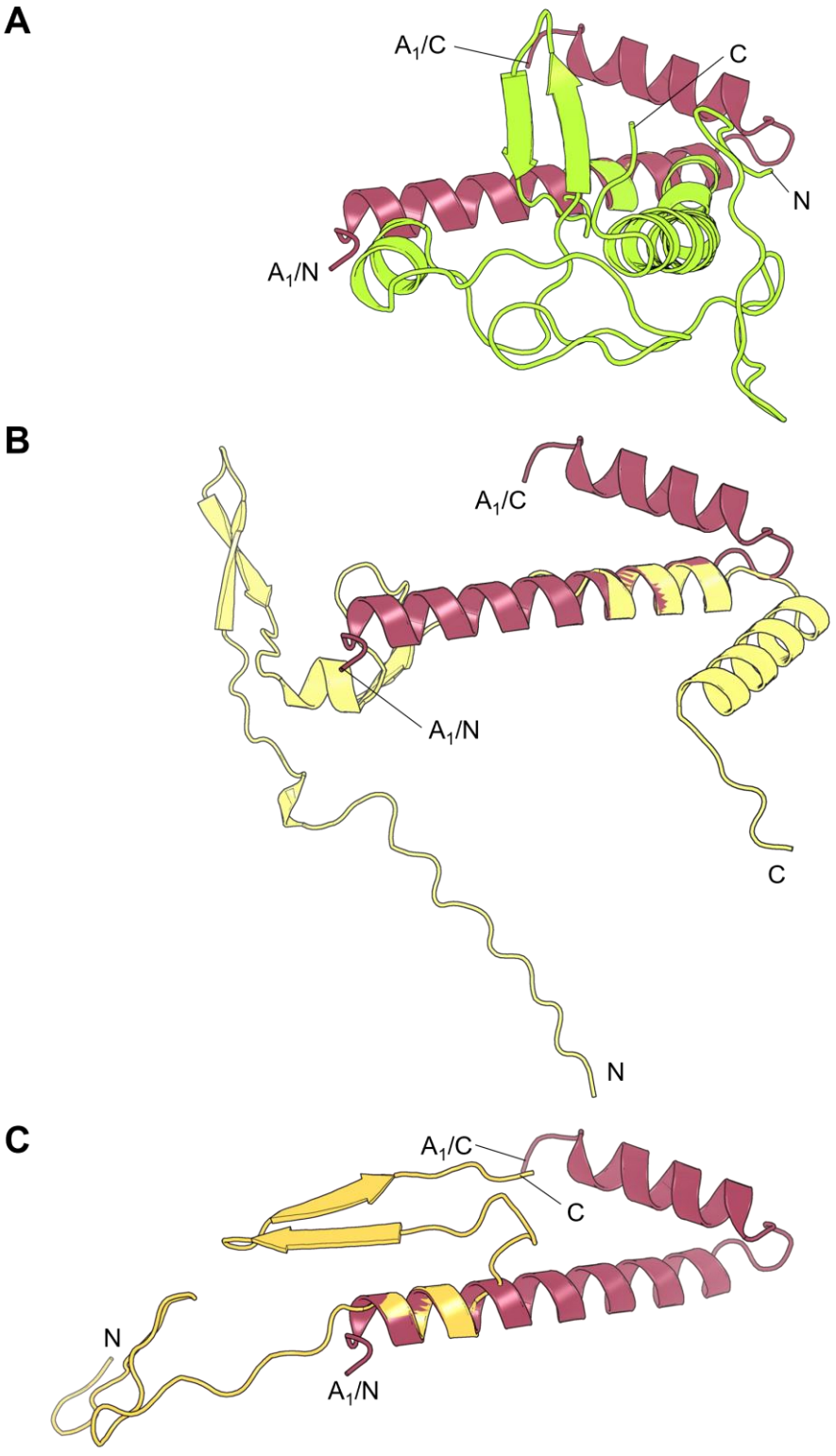


Figure 5.14. The MenA$_1$ antitoxin is structurally distinct from the type I SymE toxin. Alignment of MenA$_1$ (raspberry) to SymE models generated by **(A)** Phyre2, **(B)** AlphaFold, and **(C)** SWISS-MODEL, coloured lemon, pale yellow, and yellow orange, respectively. The alignments are displayed as cartoon representations.

5.5.4. Identification of important residues in the $MenA_1$:$MenT_1$ binding interface

The $MenA_1$:$MenT_1$ binding interface was analysed to identify key residues for protein complex formation. Close-up views revealed three distinct hydrophobic pockets to which $MenA_1$ interacts (Figure 5.15). The N-terminus L12 of $MenA_1$ interacts with 2 residues of the $MenT_1$[A] N-terminus α1 helix, which forms a hydrophobic region with $MenT_1$[A] V95 and A93, part of a linker loop connecting the NTD anti-parallel β-sheet (Figure 5.15A). Meanwhile, $MenA_1$ V19 and L14 each interact with identical hydrophobic pockets in $MenT_1$[A] and $MenT_1$[B], respectively, on either side of the $MenA_1$ central helix (Figure 5.15B). Notably, the helix and loop from $MenT_1$ F28 to F38, which feature in the identical hydrophobic binding pockets (Figure 5.15B), take a different route in each protomer, which is highlighted by aligning $MenT_1$[B] to $MenT_1$[A] (Figure 5.16A). A close-up view of the interaction interface reveals that $MenT_1$[B] F38 is pushed out by the asymmetry of $MenA_1$ binding, where it encounters $MenA_1$ M16 and Y17 (Figure 5.16B). This forms a hydrophobic network of $MenT_1$ phenylalanine residues around $MenA_1$, and indicates the likely importance of $MenT_1$ F28 and F38 to $MenA_1$ binding.

The $MenA_1$:$MenT_1$ crystal structure was next submitted to the online PDBsum protein analysis database[387], which generates a comprehensive overview of protein-protein interactions. PDBsum calculated a buried surface area (BSA) of 895 $Å^2$ between $MenA_1$ and $MenT_1$[A]. This was supported by a salt bridge formed between $MenA_1$ R27 and $MenT_1$[A] E35, and hydrogen bonds between several $MenA_1$ and $MenT_1$[A] residues (Figure 5.17). The rest of the calculated $MenA_1$:$MenT_1$[A] interface was formed through van der Waals interactions (Figure 5.17). For $MenA_1$:$MenT_1$[B], a BSA of 1189 $Å^2$ was calculated. The majority of the interface was again proposed to form through van der Waals interactions, with hydrogen bonds also formed between several $MenA_1$ and $MenT_1$[B] residues (Figure 5.18). These hydrogen bonds were distinct to those formed between $MenA_1$ and $MenT_1$[A], highlighting the asymmetrical binding of $MenA_1$ to the $MenT_1$ protomers.

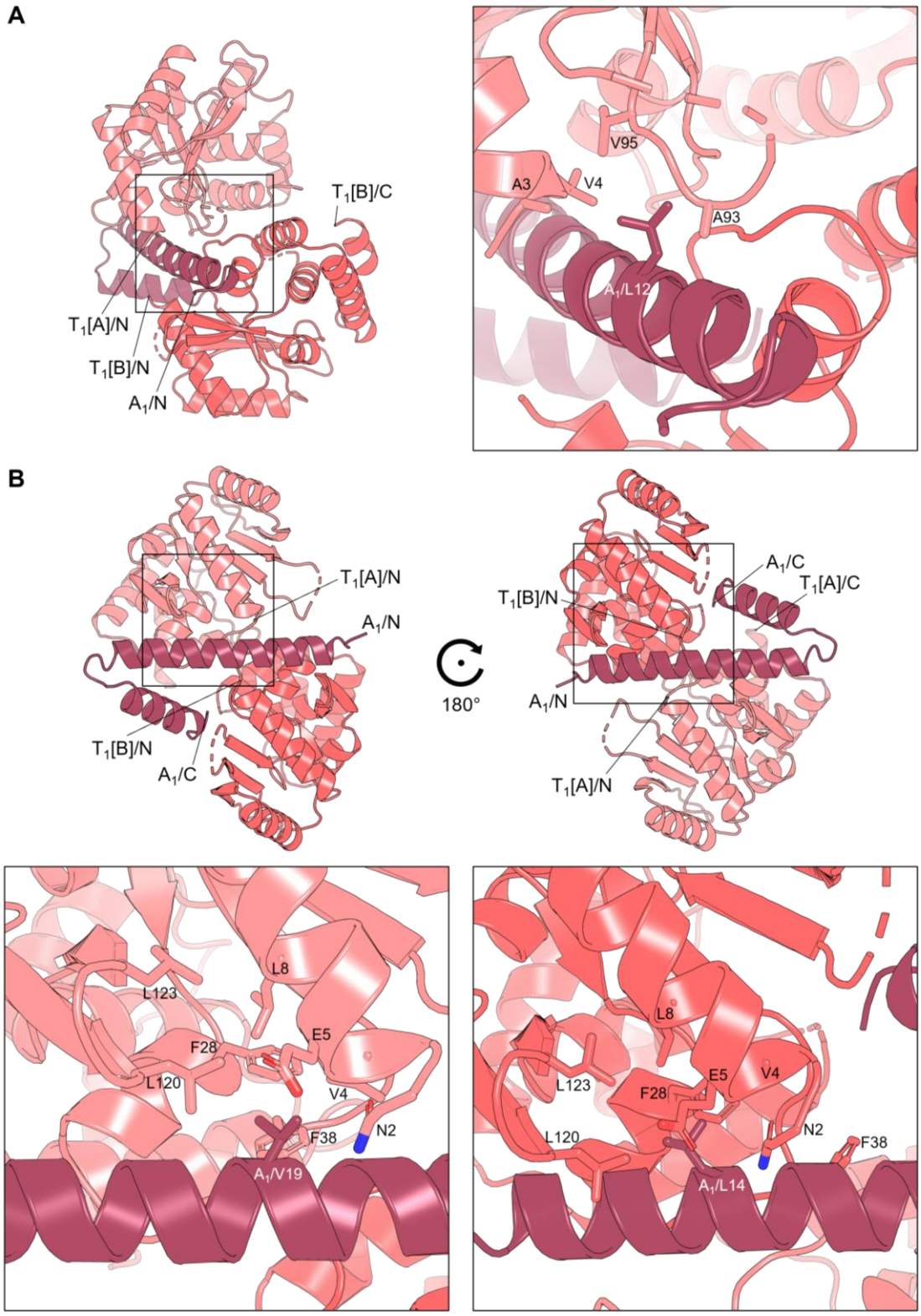


Figure 5.15. MenA$_1$ residues interact with distinct hydrophobic pockets in MenT$_1$[A] and MenT$_1$[B]. MenA$_1$ is coloured raspberry, MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. **(A)** View of the MenA$_1$ N-terminus interacting with hydrophobic residues in MenT$_1$[A]. **(B)** MenA$_1$ interacts with the same hydrophobic pocket in MenT$_1$[A] (left, top and bottom) and MenT$_1$[B] (right, top and bottom), either side of the central helix. Key residues are labelled and displayed as sticks.

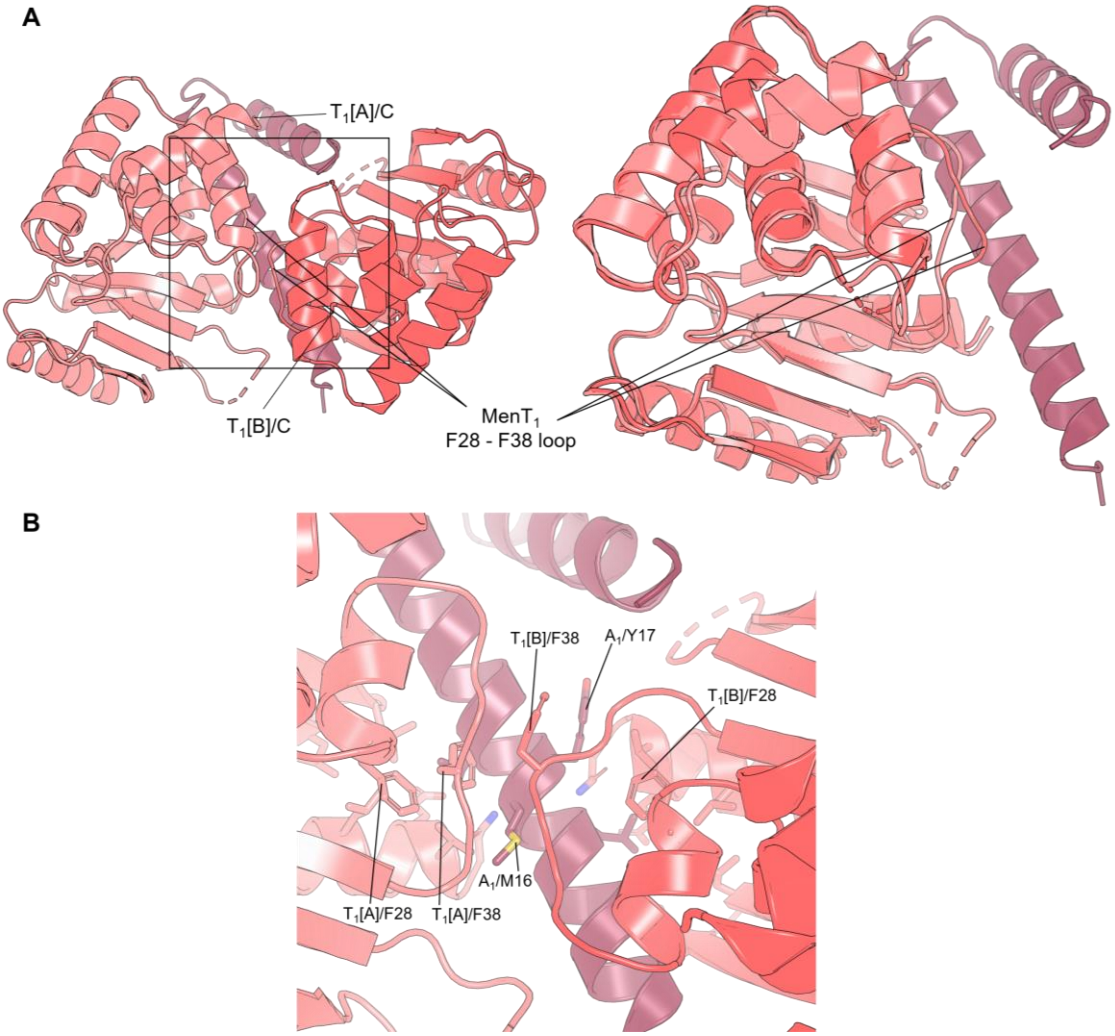


Figure 5.16. A distinct hydrophobic network is formed by asymmetrical MenA$_1$ binding. MenA$_1$ is coloured raspberry, MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. **(A)** Left: view of the MenA$_1$:MenT$_1$ crystal structure. Right: view of MenT$_1$[B] aligned to MenT$_1$[A], as in (A, left), rotated on *y* by -20° (RMSD = 0.494 Å, between 1988 atoms). **(B)** Close up view of the MenT$_1$[A] and MenT$_1$[B] F28 to F38 helix/loops interacting with the MenA$_1$ central helix, as displayed by the boxed region in (A, left). Key residues are labelled and displayed as sticks.

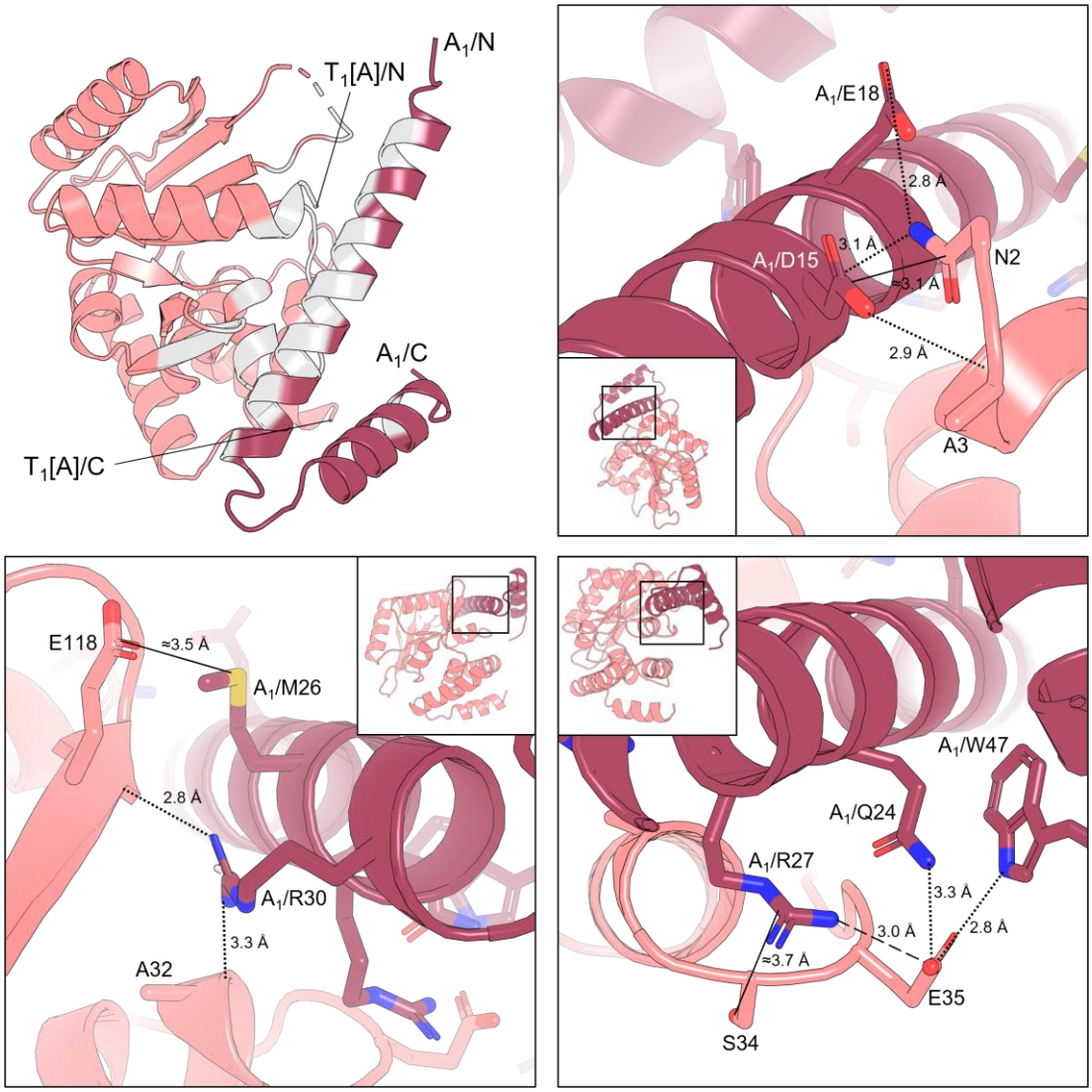


Figure 5.17. Analysis of the MenA$_1$:MenT$_1$[A] binding interface. The MenA$_1$:MenT$_1$ structure was inputted into PDBsum to investigate important residues for MenA$_1$ and MenT$_1$ protomer A ([A]) interactions. MenA$_1$ is coloured raspberry, MenT$_1$[A] is coloured salmon; the structure is displayed as a cartoon representation. Light grey residues (top, left) represent all bonded and non-bonded contacts comprising the MenA$_1$:MenT$_1$[A] binding interface. Close-up views from different angles are provided in the accompanying images. Key interface residues are displayed as sticks and the interaction types are denoted by lines: dashed line = salt bridge; dotted line = hydrogen bond; solid line = van der Waals interactions representing multiple atom interactions.

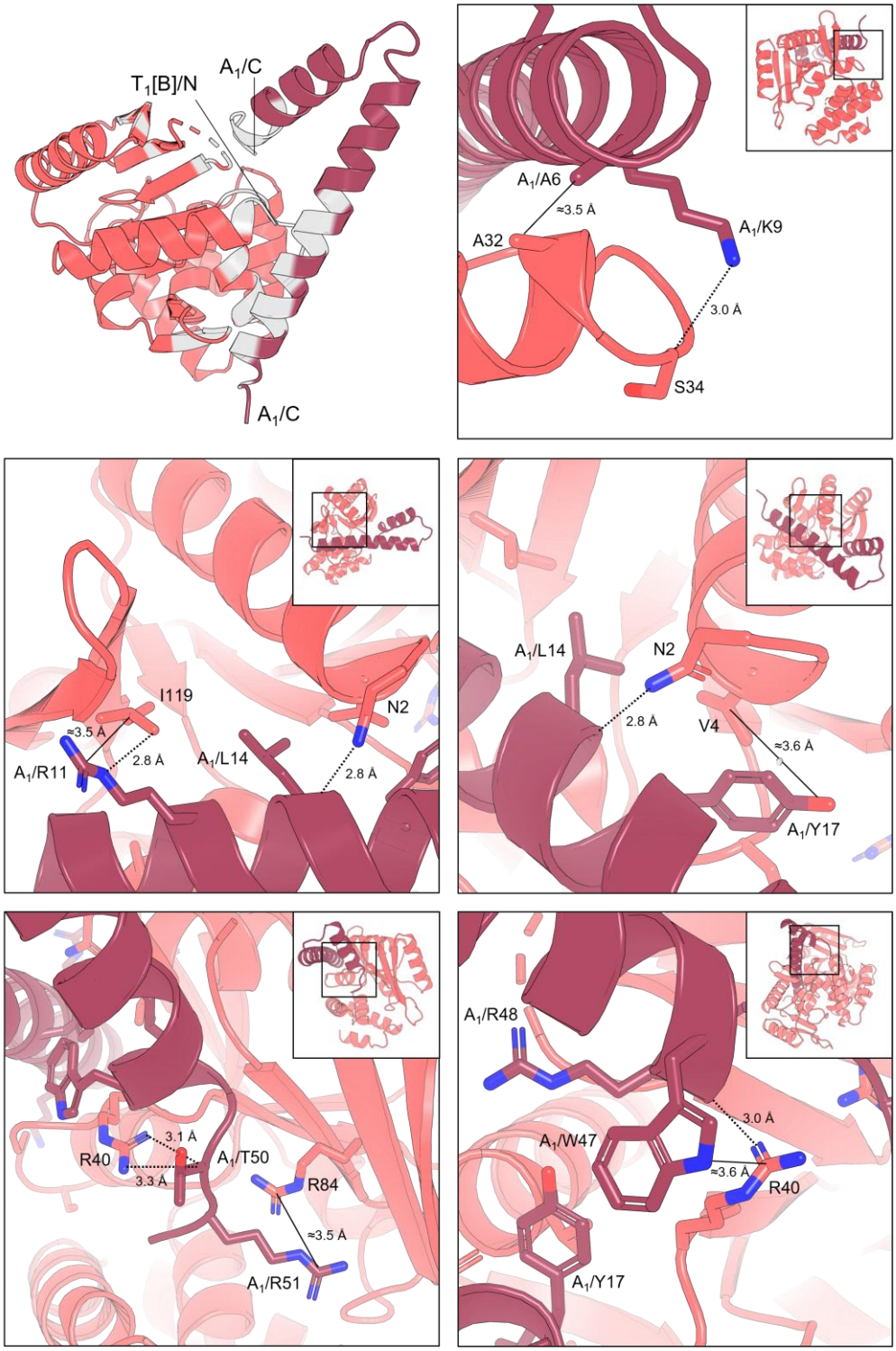


Figure 5.18. Analysis of the MenA$_1$:MenT$_1$[B] binding interface. The MenA$_1$:MenT$_1$ structure was inputted into PDBsum to investigate important residues for MenA$_1$ and MenT$_1$ protomer B ([B]) interactions. MenA$_1$ is coloured raspberry, MenT$_1$[B] is coloured deep salmon; the structure is displayed as a cartoon representation. Light grey residues (top, left) represent all bonded and non-bonded contacts comprising the MenA$_1$:MenT$_1$[B] binding interface. Close-up views from different angles are provided in the accompanying images. Key interface residues are displayed as sticks and the interaction types are denoted by lines: dotted line = hydrogen bond; solid line = van der Waals interactions representing multiple atom interactions.

5.5.5. Analysis of the $MenT_1$:$MenT_1$ protomer interface suggests the $MenT_1$ crystal structure apparent dimer is a crystallisation artifact

The $MenT_1$ protomer-protomer interface was next compared between the $MenT_1$ and $MenA_1$:$MenT_1$ crystal structures. Aligning the $MenT_1$ dimers from each structure generated a poor RMSD of 8.038 Å, between 5805 atoms (Figure 5.19A). A closer view of the aligned binding interfaces shows that the $MenT_1$[A] and $MenT_1$[B] protomers from the $MenT_1$ crystal structure are notably displaced compared to the corresponding protomers in the $MenA_1$:$MenT_1$ crystal structure (Figure 5.19A). This altered conformation causes the displacement of important $MenA_1$-binding residues, for example disruption to the $MenT_1$ hydrophobic pockets to which $MenA_1$ L14 and V19 interact (Figure 5.19B), suggesting that the $MenT_1$ crystal structure dimer would be incapable of binding $MenA_1$.

Each crystal structure was independently submitted to PDBsum[387], to generate a detailed analysis of the respective $MenT_1$ protomer-protomer binding interfaces. PDBsum calculated a $MenT_1$[A]:$MenT_1$[B] BSA of 1134 $Å^2$ and 496 $Å^2$ for the $MenT_1$ and $MenA_1$:$MenT_1$ crystal structures, respectively. Both interfaces were calculated to consist primarily of van der Waals interactions. For the $MenT_1$[A]:$MenT_1$[B] interface from the $MenA_1$:$MenT_1$ crystal structure, the importance of $MenA_1$ is apparent. $MenT_1$[A]:$MenT_1$[B] interactions were calculated primarily between residues that either form a tight network around $MenA_1$, with which they also interact, or else are between secondary structural elements that are positioned as a result of $MenA_1$ binding (Figure 5.20). This provides a biologically relevant context for the low BSA of $MenT_1$[A]:$MenT_1$[B] (496 $Å^2$) compared to $MenA_1$:$MenT_1$[A] (895 $Å^2$) and $MenA_1$:$MenT_1$[B] (1189 $Å^2$), suggesting that complex formation is dependent primarily on extensive $MenA_1$ interactions (Figure 5.20). Notably, in addition to mediating binding to $MenA_1$, $MenT_1$ F38 is also involved in $MenT_1$ protomer-protomer interactions, further highlighting its potential importance to $MenA_1$:$MenT_1$ complex formation (Figure 5.20).

In contrast, the main residues involved in the $MenT_1$[A]:$MenT_1$[B] interface from the $MenT_1$ crystal structure are more widespread, located predominantly on more flexible secondary structure loop elements (Figure 5.21). The greater distribution of these interactions reflects the displacement of the two $MenT_1$ protomers compared to the $MenA_1$:$MenT_1$ crystal structure. This further emphasises the incompatibility of $MenA_1$ binding within this complex, and suggests that the high BSA within the $MenT_1$ crystal structure apparent dimer is likely an artifact of the crystallisation process.

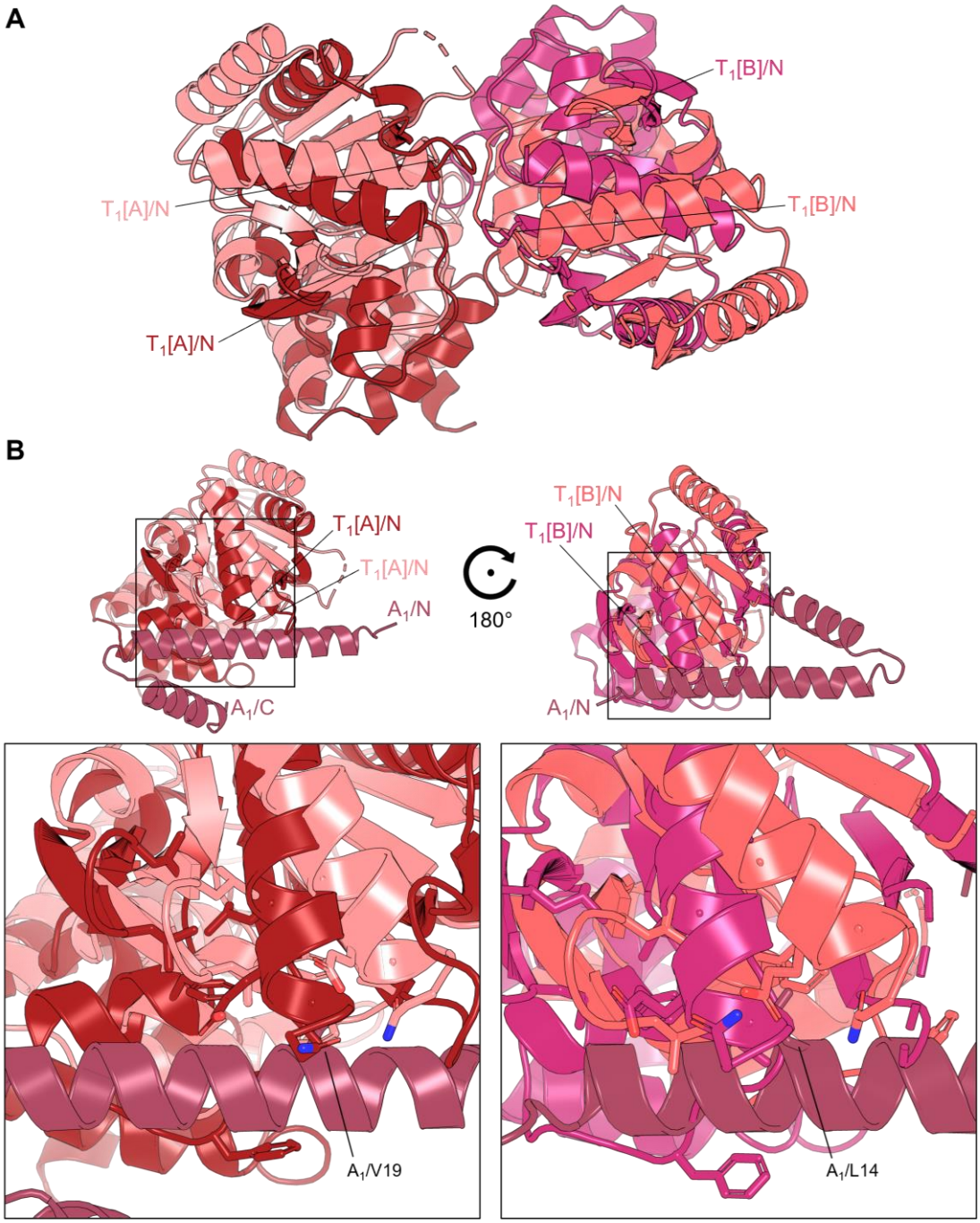


Figure 5.19. Secondary structure positioning of the MenT$_1$ crystal structure protomers suggests an inability to bind MenA$_1$. (A) Alignment of the MenT$_1$ dimers from the MenT$_1$ and MenA$_1$:MenT$_1$ crystal structures: MenT$_1$ protomer A ([A]) is coloured fire brick and salmon, respectively; MenT$_1$ protomer B ([B]) is coloured warm pink and deep salmon, respectively. The alignment is displayed as a cartoon representation. **(B)** Alignment of MenT$_1$[A] protomers (top, left) and MenT$_1$[B] protomers (top, right) from the MenT$_1$ and MenA$_1$:MenT$_1$ crystal structures, with MenA$_1$ (raspberry) superimposed into position according to MenA$_1$:MenT$_1$ crystal structure coordinates. Close-up views (bottom) show the hydrophobic pocket identified in Figure 5.15B for each protomer alignment. Key residues are displayed as sticks; MenA$_1$ residues V19 (bottom, left) and L14 (bottom, right) are labelled.

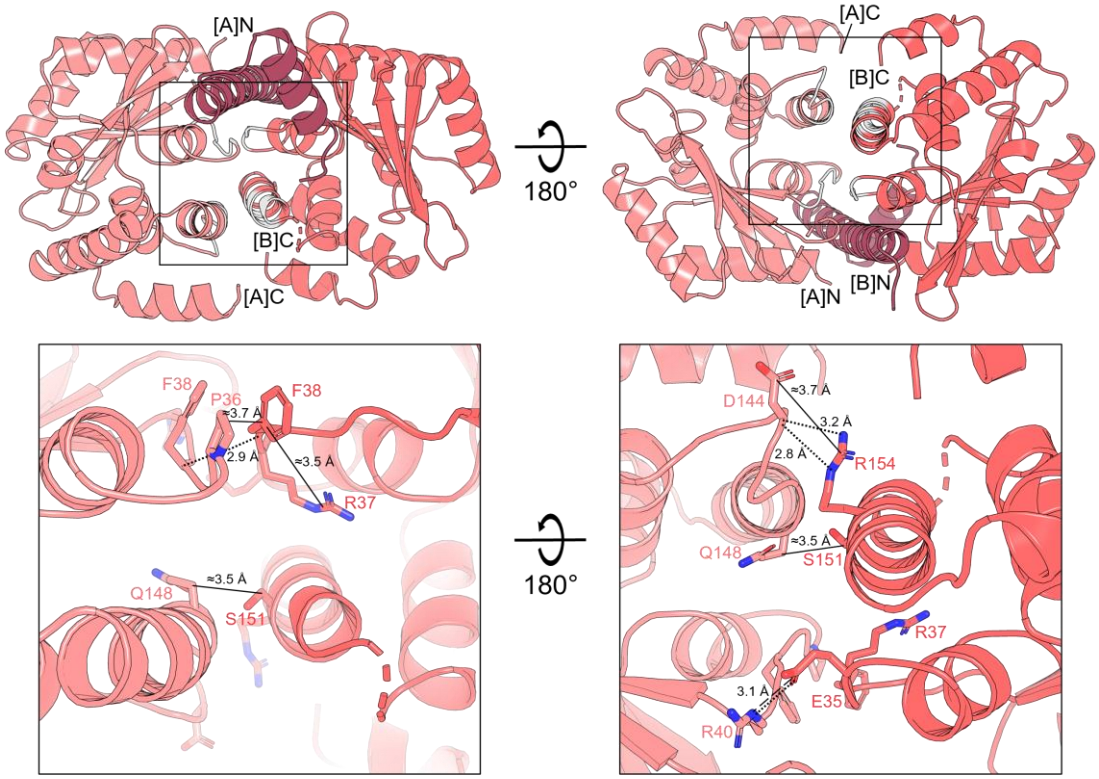


Figure 5.20. Analysis of the MenT$_1$:MenT$_1$ binding interface from the MenA$_1$:MenT$_1$ crystal structure. The MenA$_1$:MenT$_1$ crystal structure was inputted into PDBsum to identify important residues for MenT$_1$ protomer-protomer interactions. MenT$_1$ protomer A ([A]) is coloured salmon, MenT$_1$ protomer B ([B]) is coloured deep salmon; the structure is displayed as a cartoon representation. Light grey residues (top) represent all bonded and non-bonded contacts comprising the MenT$_1$[A]:MenT$_1$[B] binding interface. Close-up views (bottom) show key interface residues represented as sticks, with the interaction types denoted by lines: dashed line = salt bridge; dotted line = hydrogen bond; solid line = van der Waals interactions representing multiple atom interactions.

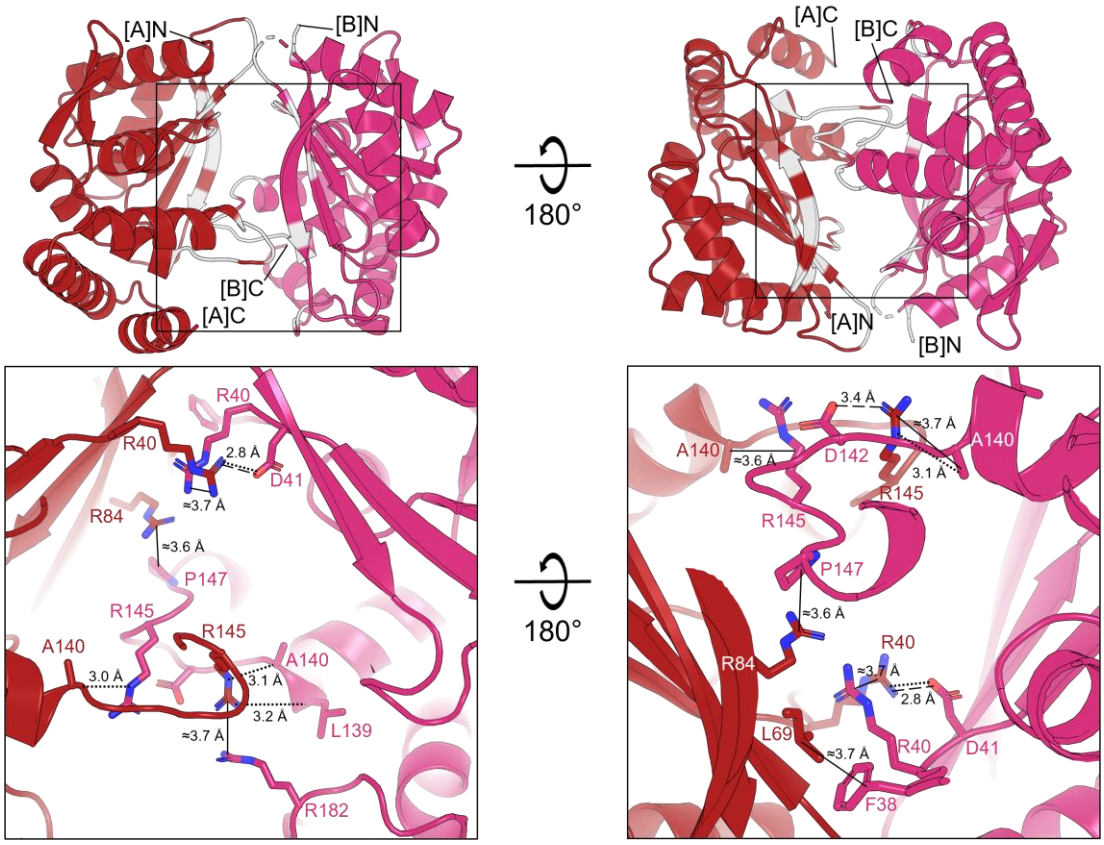


Figure 5.21. Analysis of the MenT$_1$:MenT$_1$ binding interface from the MenT$_1$ crystal structure. The MenT$_1$ crystal structure was inputted into PDBsum to identify important residues for MenT$_1$ protomer-protomer interactions. MenT$_1$ protomer A ([A]) is coloured fire brick, MenT$_1$ protomer B ([B]) is coloured warm pink; the structure is displayed as a cartoon representation. Light grey residues (top) represent all bonded and non-bonded contacts comprising the MenT$_1$[A]:MenT$_1$[B] binding interface. Close-up views (bottom) show key interface residues represented as sticks, with the interaction types denoted by lines: dashed line = salt bridge; dotted line = hydrogen bond; solid line = van der Waals interactions representing multiple atom interactions.

5.6. The MenA-MenT family consists of diverse TA classes

The five MenT crystal structures provided compelling evidence that MenA-MenT family TA systems do not belong to one overarching TA class. $MenA_1$ directly interacts and complexes with $MenT_1$ *in vitro*, as confirmed by the $MenA_1$:$MenT_1$ crystal structure, which is reminiscent of type II systems. In comparison, co-expression of $MenA_3$ and $MenT_3$ does not result in a corresponding $MenA_3$:$MenT_3$ crystal structure, indicating that these proteins do not interact, and suggestive of a type IV mechanism. Adding further complexity, $MenA_3$ co-expression instead appeared to cause phosphorylation of $MenT_3$ S78, suggesting a possible enzymatic kinase-based mode of action for $MenA_3$. In addition, as $MenA_4$ is a homologue of $MenA_3$, this suggests that $MenT_4$ toxicity is controlled as observed for $MenT_3$, potentially via a phosphorylation-based mechanism. Finally, the $MenA_4$ antitoxin has also been shown to transcriptionally regulate $MenA_4$-$MenT_4^{380}$ (See, *4.3*), further complicating its classification. As a result, $MenA_1$-$MenT_1$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$ were next investigated by analytical SEC to better characterise their respective canonical TA class. The principle of SEC, where molecules in solution are applied to a resin-packed SEC column and separated based on size, lends itself to visualising protein-protein interactions; larger molecules, such as protein complexes, are unable to fully enter the porous resin and therefore pass through the column quicker, whereas smaller molecules, such as monomeric proteins, infiltrate the resin for longer, leading to a delayed elution.

To assess $MenA_1$-$MenT_1$ interactions, the SEC profiles of previous $MenA_1$ and $MenT_1$ protein purifications were compared and analysed. The $MenA_1$:$MenT_1$ crystal structure was solved by co-incubating $MenA_1$ and $MenT_1$ proteins which had been separately expressed and purified, then setting crystallisation trials using the resulting complex (See, *5.5*). The SEC chromatogram from the $MenT_1$ purification showed that the protein eluted at a volume of 63.4 ml, with the corresponding SDS-PAGE gel revealing that this peak contained only purified $MenT_1$ protein (Figure 5.22). In comparison, the SEC chromatogram of co-incubated $MenA_1$ and $MenT_1$ showed protein eluting primarily at 52.7 ml, a shift of almost 11 ml, and indicative of a larger protein complex eluting sooner (Figure 5.22A). Analysis of the corresponding protein fractions by SDS-PAGE revealed the presence of both $MenA_1$ and $MenT_1$ proteins (Figure 5.22B), corroborating the findings in the resulting crystal structure, and confirming that $MenA_1$ and $MenT_1$ complex *in vitro*. A smaller SEC peak of approximately 65 ml was also visible after $MenA_1$-$MenT_1$ co-incubation, indicating the presence of unbound $MenT_1$ (Figure 5.22A).

To next test whether $MenA_1$ and $MenT_1$ interact *in vivo*, untagged *menA₁* was first amplified by PCR from *M*. *tuberculosis* H37Rv genomic DNA and LIC cloned into pTRB549 to generate pTRB597. Plasmid pTRB597 was then used to co-transform *E. coli* ER2566 alongside pTRB629; the resulting strain was used to co-express and co-purify $MenA_1$ and $MenT_1$ proteins (Materials and Methods, *2.6* and *2.7*). The SEC elution peak again shifted approximately 11 ml, which coupled with SDS-PAGE analysis showing the presence of both $MenA_1$ and $MenT_1$ proteins, also demonstrated that $MenA_1$ and $MenT_1$ interact *in vivo*. The fact that $MenA_1$:$MenT_1$ elutes at the same volume following either co-expression or co-incubation indicates that the complex in solution matches that of the crystal structure: i.e., monomeric $MenA_1$ and dimeric $MenT_1$, a mass of 50.73 kDa. The elution profile of $MenT_1$ on the other hand, almost 11 ml later than that of $MenA_1$:$MenT_1$, does not appear to correspond to that of a 43.21 kDa $MenT_1$ dimer, as comprises the $MenT_1$ crystal structure. When coupled with the $MenT_1$ binding interface data, which suggests that the dimeric $MenT_1$ structure was likely a crystallisation artifact (See, *5.5.5*), the SEC elution profile suggests that $MenT_1$ is instead a monomer in solution.

A similar analysis was performed on the SEC profiles from $MenT_3$ and $MenA_3$+$MenT_3$ expressions (See, *5.3*), to explore for possible $MenA_3$:$MenT_3$ interactions *in vivo* (Figure 5.23A). The $MenT_3$ SEC elution profile was comparable for $MenT_3$ expressed alone, or co-expressed alongside $MenA_3$, indicating that the antitoxin and toxin do not interact (Figure 5.23A). This was confirmed by SDS-PAGE analysis, where no $MenA_3$ protein was detectable in the $MenA_3$+$MenT_3$ co-expression sample (Figure 5.23B). Next, $MenA_3$ was separately expressed and purified to first confirm whether it can be produced, and then to provide a control for comparison to the $MenA_3$+$MenT_3$ co-expression. The *menA₃* ORF was first amplified by PCR from pPF656, then LIC cloned into pSAT1-LIC to produce pTRB491, encoding an N-terminal His_6-SUMO-$MenA_3$ fusion, which was then used to transform *E. coli* ER2566. This strain was then used to express and purify $MenA_3$ (Materials and Methods, *2.6* and *2.7*). Interestingly, $MenA_3$ was particularly unstable compared to the other MenA/MenT proteins purified, often precipitating out of solution in standard purification buffers, and following buffer optimisation, seemed more stable in high NaCl/KCl concentrations of >0.8 M. As a result of its relative instability, the number of $MenA_3$ purification steps were reduced, with SEC omitted entirely due to fears of protein precipitation and consequently low yields that would affect downstream experiments. Instead, SDS-PAGE analysis was performed on $MenA_3$ protein samples collected after the second affinity chromatography step, which confirmed the presence of the antitoxin (Figure 5.23B).

Having shown that $MenA_3$-$MenT_3$ did not appear to interact *in vivo*, analytical SEC was next performed to confirm the expected absence of *in vitro* interactions. Purified $MenA_3$ antitoxin and $MenT_3$ toxin proteins from separate expressions were applied to a high-resolution SEC column, either independently or following co-incubation of the cognate proteins (Materials and Methods, *2.10*) (Figure 5.23C). Analytical SEC chromatograms confirmed no shift in the elution profile of $MenT_3$ when mixed with $MenA_3$, suggesting that these systems exhibit no interactivity *in vitro* (Figure 5.23C). Interestingly, the 22.40 kDa $MenA_3$ antitoxin eluted sooner than the 31.86 kDa $MenT_3$ toxin when tested independently (Figure 5.23C). Given that the $MenA_4$ and AbiEi homologues are both monomeric[84,378], this likely reflects the inherent instability of $MenA_3$ and possibly dysfunctional $MenA_3$ protein aggregation during SEC.

The *in vivo* interaction profiles of $MenA_4$ and $MenT_4$ were also assessed. Firstly, *$menA_4$* was amplified by PCR from pPF658 and LIC cloned into pSAT1-LIC to produce pTRB493. *E. coli* ER2566 was then separately transformed with either pTRB493 (*His_6-SUMO-$menA_4$*), pTRB544 (*His_6-SUMO-$menT_4$*), or pPF658 (*$menA_4$*) and pTRB544, then grown and used to express and purify the respective $MenA_4$/$MenT_4$ combinations (Materials and Methods, *2.6* and *2.7*). The protein yields from the $MenA_4$+$MenT_4$ co-expression were notably poor; after AEC purification, the chromatogram peak representing the relative protein concentration was barely detectable (measured as UV absorbance at 280 nm). The reasons for such poor yields were unknown as the protein yield derived from the $MenT_4$-only expression was comparably higher, suggesting that $MenT_4$ toxicity was not a factor. During the *E. coli* antitoxicity assays described previously (See, *3.4*; Figure 3.7), co-expression of $MenA_4$ alongside $MenT_4$ was able to neutralise $MenT_4$ toxicity and restore viable counts; however, bacterial colonies were visibly smaller and sicker looking than when $MenA_4$ and $MenT_4$ expression was uninduced. This might indicate that the co-expression of $MenA_4$ and $MenT_4$ exerts an uncharacterised cooperative inhibitory effect on cell growth or metabolism, or potentially exacts a fitness cost which is reflected in the low protein yield from a $MenA_4$+$MenT_4$ expression.

Purification of the $MenA_4$+$MenT_4$ expression sample was therefore stopped after the second affinity chromatography step; the protein yield at this stage was deemed too low to proceed with extra purification, as each additional step typically results in small but noticeable protein loss. As a result, the purified samples were analysed only by SDS-PAGE (Figure 5.24A). These showed no detectable $MenA_4$ in the co-expression sample, indicating that $MenA_4$ and $MenT_4$ likely do not interact *in vivo*. To test whether $MenA_4$ and $MenT_4$ interact *in vitro*, separately expressed $MenA_4$ and $MenT_4$ proteins were applied to the high-resolution SEC column either

independently or following a co-incubation (Materials and Methods, *2.10*) (Figure 5.24B). As observed with $MenA_3$-$MenT_3$, the analytical SEC chromatograms showed no change in the $MenT_4$ elution profile, regardless of the presence of $MenA_4$, suggesting that this system also does not interact *in vitro* (Figure 5.24B).

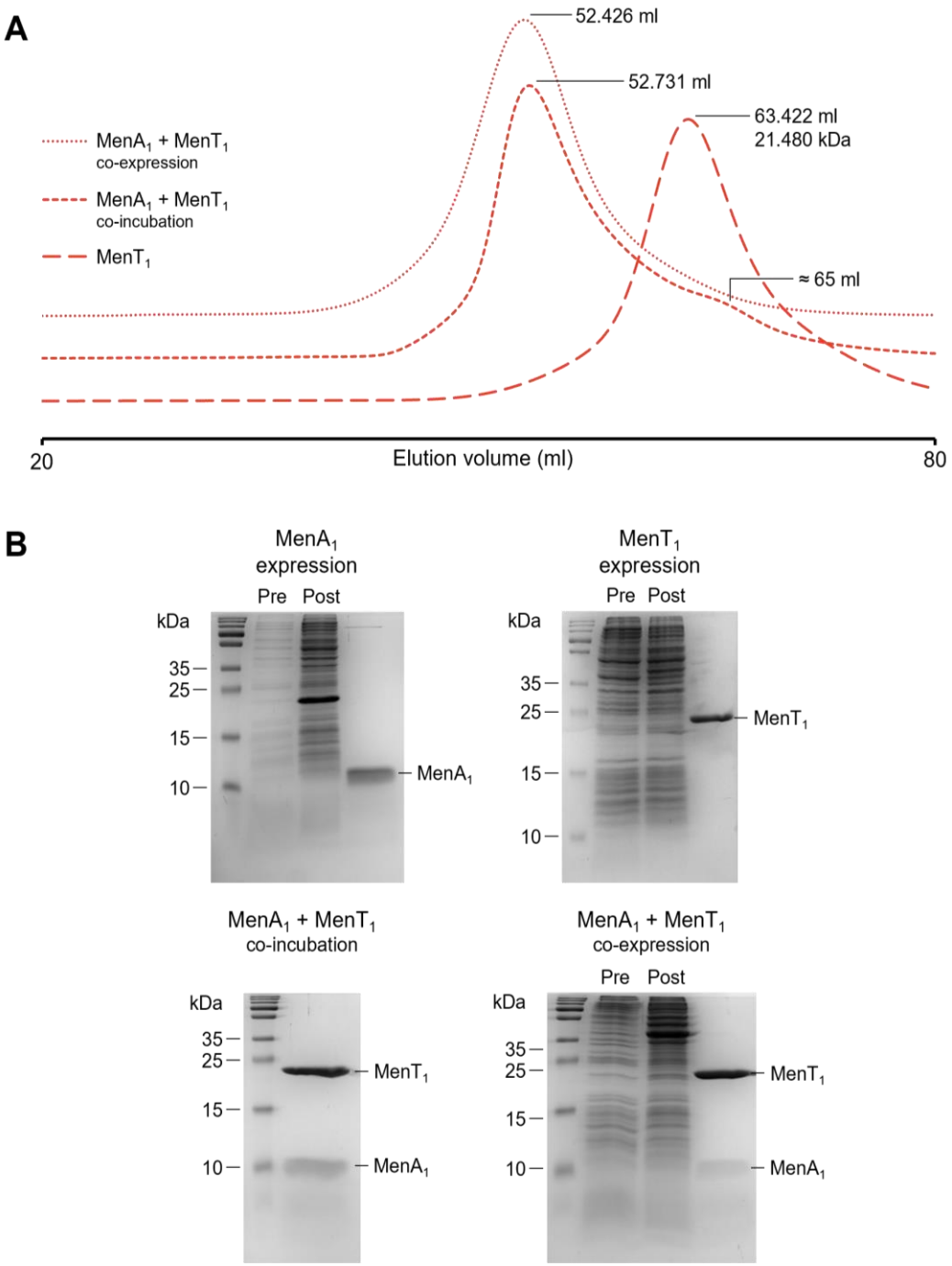


Figure 5.22. $MenA_1$-$MenT_1$ interact *in vitro* and *in vivo*. (A) SEC chromatograms of $MenT_1$ expressed alone, co-expressed with $MenA_1$ *in vivo*, or expressed and purified separately, then co-incubated *in vitro*. Relative protein concentration was measured as UV absorbance at 280 nm. **(B)** SDS-PAGE analysis of the purified $MenA_1$ antitoxin, in addition to SDS-PAGE analysis of the proteins purified by SEC in (A). The toxin-only expression sample was confirmed by MS.

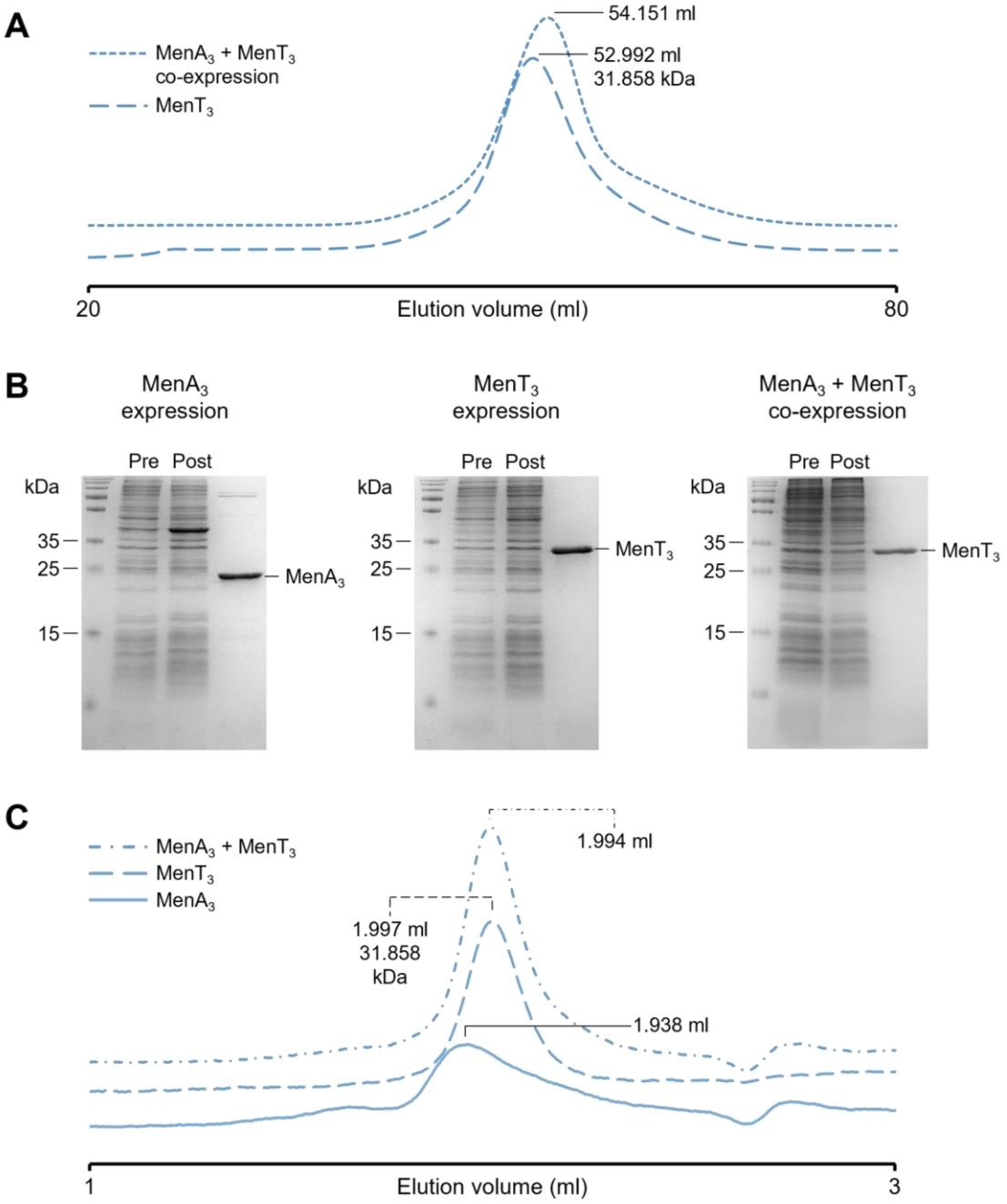


Figure 5.23. MenA₃-MenT₃ do not interact. (A) SEC chromatograms of MenT₃ expressed alone or co-expressed with MenA₃ *in vivo*. Relative protein concentration was measured as UV absorbance at 280 nm. **(B)** SDS-PAGE analysis of purified MenA₃ antitoxin protein, and the MenT₃ toxins purified by SEC in (A). The toxin-only expression sample was verified by MS. **(C)** MenA₃ and MenT₃ were expressed and purified separately, then assessed by high resolution analytical SEC (Materials and Methods, *2.10*). To test toxin-antitoxin interactions and complex formation *in vitro*, toxins and antitoxins were first mixed at a 1:1 molar ratio and incubated for 30 min at room temperature, then applied to the analytical SEC column. Relative protein concentration was measured as UV absorbance at 280 nm. All samples were tested in triplicate. The presented data for each sample is representative of all corresponding replicates.

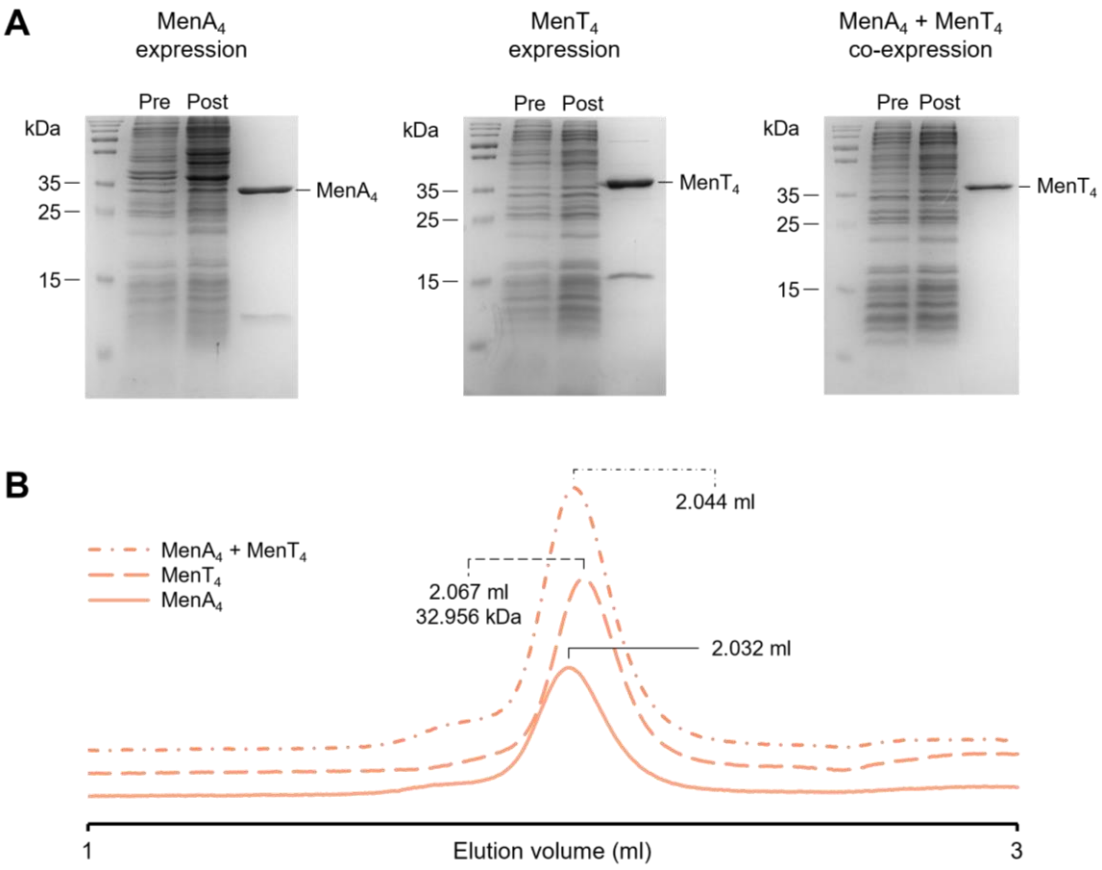


Figure 5.24. MenA4-MenT4 do not interact. (A) SDS-PAGE analysis of purified MenA4 protein (post-SEC) from lone antitoxin expression, purified MenT4 (post-SEC) from lone toxin expression, and purified MenT4 (post-AEC) from an *in vivo* co-expression with MenA4. **(B)** MenA4 and MenT4 were expressed and purified separately, then assessed by high resolution analytical SEC (Materials and Methods, *2.10*). To test toxin-antitoxin interactions and complex formation *in vitro*, toxins and antitoxins were first mixed at a 1:1 molar ratio and incubated for 30 min at room temperature, then applied to the analytical SEC column. Relative protein concentration was measured as UV absorbance at 280 nm. All samples were tested in triplicate. The presented data for each sample is representative of all corresponding replicates.

5.7. MenA-mediated post-translational modifications of MenT toxins

The characterisation of the MenA-MenT TA systems revealed that $MenA_1$ and $MenT_1$ complexed in solution, which was confirmed by the $MenA_1$:$MenT_1$ crystal structure (Figures 5.11 and 5.22), whereas $MenT_3$ and $MenT_4$ appeared not to interact with their cognate antitoxins (Figures 5.23 and 5.24). In the case of $MenA_1$-$MenT_1$, the interaction between antitoxin and toxin served to explain the mechanism of antitoxicity, where $MenA_1$ bridges and aids assembly of a $MenT_1$:$MenA_1$:$MenT_1$ complex, blocking the toxin active sites. However, the antitoxic mechanism for $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ remained puzzling. For example, the presence of a phosphoserine at S78 of the $MenT_3$ SEP crystal structure was putatively linked to $MenA_3$ co-expression, and a $MenT_3^{(S78A)}$ mutant was recalcitrant to $MenA_3$ antitoxicity (Figure 5.5B). This phenotype suggests that the experimentally observed $MenT_3$ S78 phosphoserine might represent an important mechanistic component of $MenA_3$ antitoxicity. Yet for $MenA_4$-$MenT_4$, mutating the equivalent conserved serine in $MenT_4$ (S67A) instead abolished toxicity (Figure 5.8A). The divergent phenotypes were puzzling, perhaps reflecting that the biological role of the S78/S67 serine is not conserved between $MenT_3$ and $MenT_4$, and therefore that these toxins are neutralised by an evolutionarily distinct mechanism. Alternatively, the lost toxicity of $MenT_4^{(S67A)}$ might simply mask its essentiality to $MenA_4$ activity, and may reflect a bifunctional role for this conserved residue in both toxicity and antitoxicity.

Consequently, to further investigate a conserved basis for antitoxicity across the MenA-MenT family and better explore the molecular mechanisms at play, mass spectrometry (MS) analyses were performed. These experiments tested MenT protein samples produced with and without co-expression of the cognate antitoxin, to detect the presence or absence of possible MenA-mediated PTMs corresponding to phosphorylation. Purified MenT protein samples were used that were previously tested in protein interaction studies (See, *5.6*; Figures 5.22 to 5.24), with the exception of the $MenA_1$+$MenT_1$ co-expression sample. For this sample, the antitoxin and toxin were shown to interact *in vitro* and *in vivo* (Figure 5.22). However, during AEC purification of the $MenA_1$+$MenT_1$ co-expression, two distinct chromatogram peaks were observed: one peak contained both $MenA_1$ and $MenT_1$ proteins, presumably complexed, as revealed by SDS-PAGE, and which was used for interaction studies; the second contained only $MenT_1$ protein, which had presumably disassociated from the $MenA_1$:$MenT_1$ complex. This second sample was independently purified as per Materials

and Methods, *2.7*, and used for subsequent MS experiments. The toxin proteins were prepared for MS analyses as described in Materials and Methods, *2.11*.

The results were intriguing. The predicted mass of $MenT_1$ is 21481 Da, which corresponded to the dominant analyte in the MS sample derived from lone $MenT_1$ expression (Figure 5.25A). In the corresponding $MenA_1$+$MenT_1$ co-expression sample, the dominant analyte was also 21480 Da (Figure 5.25B). However, a second species was also observed, approximately 50% of the relative concentration of the dominant analyte, with a mass of 21560 Da (Figure 5.25B). This 80 Da shift potentially corresponds to an AA phosphorylation modification, which was surprising, given that the $MenA_1$:$MenT_1$ crystal structure and SEC results suggested that formation of the T:A complex would explain the mechanism of antitoxicity, and there was no evidence of phosphorylation in the crystal structure.

MS analysis of $MenT_3$ protein samples confirmed the proposed link between $MenA_3$ co-expression and phosphorylation: when $MenT_3$ was expressed alone, the dominant analyte had a mass of 31858 Da, corresponding to the predicted $MenT_3$ mass of 31859 Da (Figure 5.26A); following $MenA_3$ co-expression, the dominant species shifted in mass by 80 Da to 31938 Da (Figure 5.26B), corresponding to the observed phosphoserine in the $MenT_3$ SEP crystal structure. Together, these observations suggest that $MenA_3$ mediates the phosphorylation of $MenT_3$ at serine S78, potentially through direct kinase activity. A small peak consistent with unphosphorylated $MenT_3$ was also present in the $MenA_3$+$MenT_3$ co-expression sample (31858 Da). This might reflect the constant synthesis of *de novo* $MenT_3$ during forced expression, or that putative $MenA_3$ kinase activity is potentially inefficient; alternatively, it could be the result of an active phosphatase removing the modification. In contrast, the dominant analytes in MS samples from both lone $MenT_4$ expression and $MenA_4$+$MenT_4$ co-expression closely matched the predicted $MenT_4$ mass of 32957 Da (Figure 5.27), suggesting that in the $MenA_4$-$MenT_4$ system, the antitoxin does not mediate toxin phosphorylation.

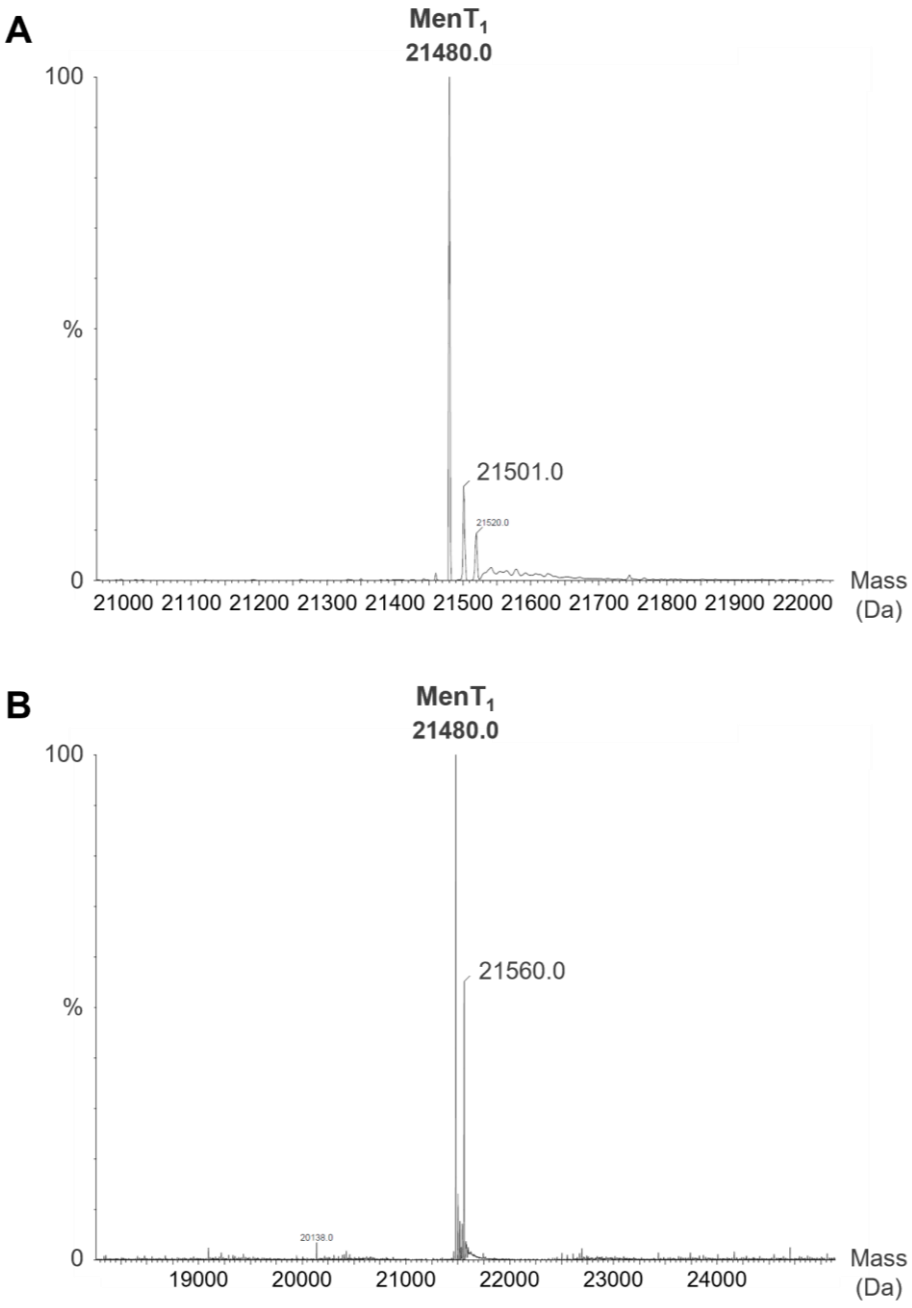


Figure 5.25. MenA$_1$ co-expression leads to a heterogeneous MenT$_1$ population. (A) ES+ TOF MS of purified MenT$_1$ toxin protein after lone expression. **(B)** ES+ TOF MS of purified MenT$_1$ following co-expression with MenA$_1$ antitoxin. Experiments were performed once; data are presented by mass, and as a relative percentage of the dominant analyte. Please note, the *x* range differs per graph.

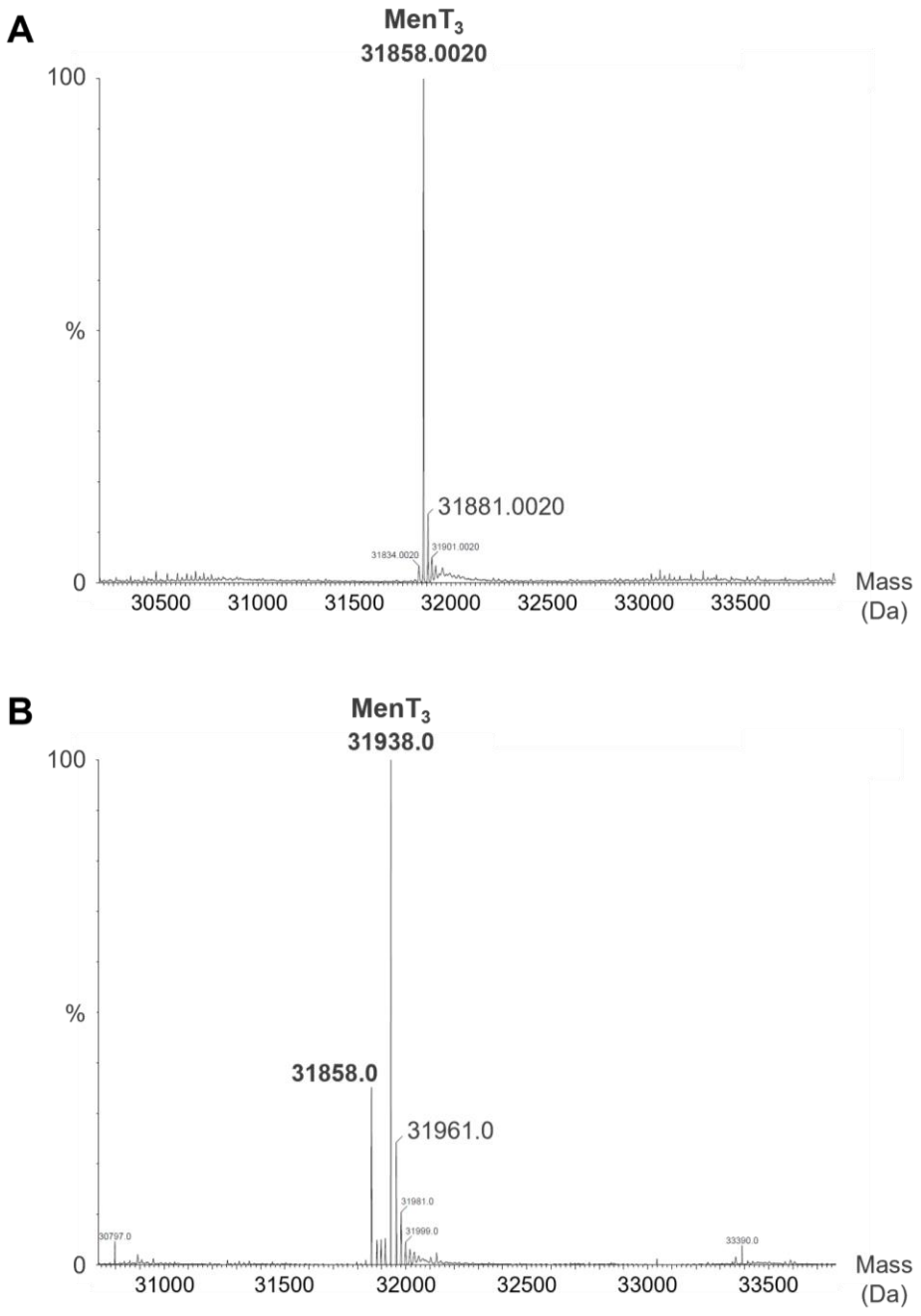


Figure 5.26. MenA$_3$ mediates MenT$_3$ phosphorylation. (A) ES+ TOF MS of purified MenT$_3$ toxin protein after lone expression. **(B)** ES+ TOF MS of purified MenT$_3$ following co-expression with MenA$_3$ antitoxin. Experiments were performed once; data are presented by mass, and as a relative percentage of the dominant analyte. Please note, the *x* range differs per graph.

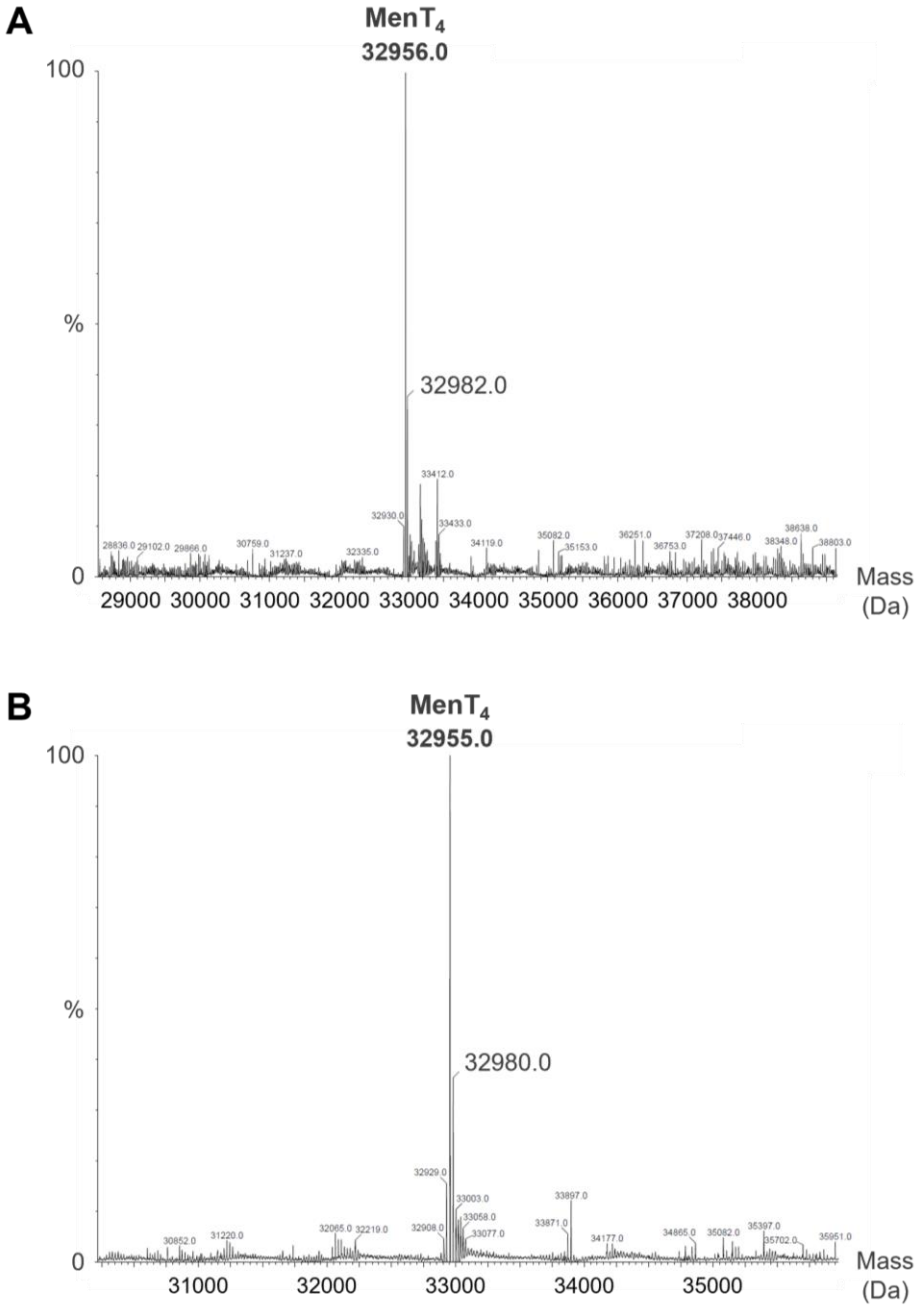


Figure 5.27. MenA$_4$ co-expression does not lead to MenT$_4$ modifications. (A) ES+ TOF MS of purified MenT$_4$ toxin protein after lone expression. **(B)** ES+ TOF MS of purified MenT$_4$ following co-expression with MenA$_4$ antitoxin. Experiments were performed once; data are presented by mass, and as a relative percentage of the dominant analyte. Please note, the *x* range differs per graph.

5.8. Discussion

The structural characterisation of the MenT toxins was performed to investigate mode of toxicity and to identify conserved features shared by these homologous proteins. X-ray crystallography was used to determine the first crystal structures of DUF1814 toxins, specifically $MenT_1$, $MenT_3$, and $MenT_4$. These showed overall structural similarity, while mutagenesis studies of conserved $MenT_3$ and $MenT_4$ residues revealed a conserved toxin active site. This active site exhibits a strong electropositive charge, which would aid the binding of negatively charged NTPs, supporting the proposed activity of the NTase-like DUF1814 toxins.

Surprising variety was observed among the MenT crystal structures: $MenT_1$ was solved as an apparent dimer within the crystals (Figure 5.9), and as a $MenT_1$:$MenA_1$:$MenT_1$ complex (Figure 5.11); $MenT_3$ was solved with and without phosphorylation at residue S78, which was dependent on the presence of $MenA_3$ during co-expression (Figures 5.2 and 5.3); whilst $MenT_4$ was shown to be a monomer with no additional modifications (Figure 5.6). The diversity among $MenT_1$ and $MenT_3$ crystal structures could be linked to the antitoxic activity of the cognate antitoxin and were further assessed by SEC protein interaction studies and MS analyses. These experiments indicated that the MenA-MenT TA family consists of a variety of canonical TA classes. $MenT_1$ interacted with $MenA_1$ *in vitro* and *in vivo*, suggesting a type II antitoxic mechanism (Figure 5.22). $MenT_4$ demonstrated no interactivity with $MenA_4$ by SEC, and remained unmodified, suggesting a type IV antitoxic mechanism (Figures 5.24 and 5.27). In contrast, whilst $MenA_3$ and $MenT_3$ did not form complexes *in vitro* or *in vivo* (Figure 5.23), $MenT_3$ was phosphorylated at residue S78 in a $MenA_3$-dependent manner (Figure 5.26). This suggests that $MenA_3$-$MenT_3$ represents a distinct and novel mechanism of antitoxicity.

The observed $MenT_3$ S78 phosphoserine, detected in the $MenT_3$ SEP crystal structure and by MS, was reliant on co-expression of $MenA_3$. Coupled with the lack of interactivity between $MenA_3$ and $MenT_3$, suggesting no direct inhibition of toxicity by T:A complexing, this phosphorylation event might represent an important mechanistic component of $MenA_3$ antitoxicity. Given that the S78 phosphoserine in question protrudes into the $MenT_3$ active site, $MenA_3$-mediated phosphorylation may serve to sterically obstruct substrate binding. Additionally, S78 phosphorylation might disrupt the net positive charge of the active site to impair interactions with negatively charged nucleotides, consistent with the predicted NTase activity of MenT toxins.

The importance of the S78 residue to antitoxicity was demonstrated by toxicity and antitoxicity assays in *E. coli* testing a $MenT_3^{(S78A)}$ substitution mutant. These results showed that $MenT_3^{(S78A)}$ was recalcitrant to $MenA_3$ antitoxicity (Figure 5.5B), suggesting that $MenA_3$ was unable to recognise and phosphorylate the S78A mutant, and thereby neutralise $MenT_3$ toxicity. Consequently, this highlights phosphorylation of $MenT_3$ S78 by $MenA_3$ as a potentially vital component of antitoxicity. Indeed, a separate study released in 2020 by Yu *et al*. reported that $MenA_3$ did indeed exhibit kinase activity against $MenT_3$, and that $MenA_3$-mediated phosphorylation of S78 comprises the antitoxic mechanism for this system[99]. These findings categorise $MenA_3$-$MenT_3$ within the recently defined type VII TA classification[94]. Type VII TA systems are characterised by an antitoxin which enzymatically modifies the cognate toxin to inactivate it. Two other systems have so far been characterised that fit these criteria: the TomB-Hha system, in which the TomB antitoxin is proposed to oxidise a conserved Hha residue to destabilise toxin structure[96]; and MntA-HepT, where MntA polyadenylates HepT residue Y104 adjacent to the toxin's RNase motif to block activity[97]. Along with the novel kinase activity of $MenA_3$, these three systems comprise a hitherto overlooked TA class which may be considerably more widespread than previously understood.

Phosphorylation is fundamental to bacterial homeostasis, where it has established roles in cell signalling, cell cycle regulation, protein localisation and molecular transport, as well as proteolysis[391–396]. It also plays a key role in bacterial virulence, pathogenicity, and the activity of stress-responsive two-component systems (TCSs)[395,397–399]. TCSs typically feature a kinase sensor protein, which detects changes to specific conditions, and a DNA-binding response regulator, which is phospho-activated by the kinase component to regulate gene expression[397,398]. Examples of these systems include the *S*. Typhimurium PhoP-PhoQ TCS, which promotes virulence and bacterial survival under Mg^{2+}-limiting conditions[400]. Given its underlying importance to essential bacterial processes, and in promoting adaptability and survivability, it is no surprise that phosphorylation has also been linked to toxin activity in several TA systems. The type II PezT toxin phosphorylates effectors of peptidoglycan synthesis to inhibit cell wall synthesis[121], whereas the type II Doc toxin phosphorylates the EF-Tu translation elongation factor to prevent tRNA-ribosome processing[68]. However, its role as a component in antitoxicity has not been reported before. The kinase activity of $MenA_3$, predicted here and confirmed elsewhere[99], therefore presents a novel mechanism of type VII antitoxicity, and a novel antitoxic mechanism among TA systems in general. Given its ubiquity

in a variety of bacterial stress responses, phosphorylation likely represents an underreported phenomenon in TA system regulation and activity.

In the context of the $MenA_1$-$MenT_1$ system, the observation by MS of a predicted phosphorylated $MenT_1$ species following $MenA_1$ co-expression is puzzling. SEC experiments suggested that $MenA_1$ and $MenT_1$ directly interact and complex *in vitro*, which is confirmed by the $MenA_1$:$MenT_1$ crystal structure. This structure strongly suggests that antitoxicity is mediated by $MenA_1$ bridging a $MenT_1$ dimer, occluding both $MenT_1$ active sites. It is important to note that the $MenA_1$:$MenT_1$ crystal structure was obtained following the *in vitro* co-incubation of separately expressed and purified $MenA_1$ and $MenT_1$. Any *in vivo* modifications mediated by $MenA_1$ during co-expression would therefore not be reflected in the crystal structure presented here (Figure 5.11).

However, given the distinct protein size and AA sequence composition of $MenA_1$ compared to the other MenA antitoxins (68 AAs compared to 207 AAs for $MenA_3$; sequence identity = 18.3%), as well as the relatively minimalist $MenA_1$ structure (two helices connected by a short, flexible loop), it is difficult to understand how it would perform kinase activity similar to $MenA_3$. Indeed, the catalytic core of typical kinases consists of a small, predominantly β-sheet N-terminal subdomain, and a larger, mainly α-helical C-terminal subdomain[401], which is considerably at odds with the $MenA_1$ structure (Figure 5.11). Furthermore, $MenA_1$ shows no structural homology to Phyre2- and AlphaFold-predicted $MenA_3$ structures[361,362] (alignments (RMSD) = 13.635 Å, between 247 atoms, and 12.144 Å, between 247 atoms, respectively; superpositions (RMSD) = 6.147 Å, between 176 atoms, and 6.050 Å, between 205 atoms, respectively), indicating that a kinase mechanism similar to $MenA_3$ is highly unlikely. No structural homology is seen to $MenA_4$ or AbiEi either[378,380] (alignments (RMSD) = 0.423 Å, between 50 atoms, and 2.560 Å, between 161 atoms, respectively; superpositions (RMSD) = 7.523 Å, between 224 atoms, and 7.744 Å, between 385 atoms, respectively), which appears to also rule out a conserved DNA-binding component of antitoxicity.

The observed 80 Da shift may reflect a different PTM, for example sulfation, however the biological cause and function of any modification remains unexplained. The fact that the "modified" analyte is not the dominant species in the sample might reflect that the predicted phosphorylation is not an essential part of antitoxicity. Perhaps phosphorylation lends a transient and reversible extra layer to antitoxicity, which is more adaptable than the proposed quasi-permanence of T:A complexing[20]. $MenT_1$ phosphorylation could be mediated by $MenA_1$ interacting with a secondary kinase, which might then signal $MenT_1$ for cellular

translocation or proteolytic degradation[396,399]. Another possibility could be that phosphorylation is important for $MenA_1$:$MenT_1$ recognition and binding, as reported for both eukaryotes and prokaryotes[402,403]. However, this does not explain why unmodified $MenT_1$, produced in the absence of $MenA_1$ co-expression, was able to complex with $MenA_1$ *in vitro* (Figures 5.11 and 5.22). In this context, the predicted phosphorylation seen by MS might enhance $MenA_1$:$MenT_1$ complex stability[403]. Alternatively, given that the purified $MenT_1$ co-expression sample used for MS appeared to have separated from the $MenA_1$:$MenT_1$ complex during AEC, *in vivo* phosphorylation may instead serve as a regulatory mechanism to destabilise the complex and promote its disassociation, freeing the toxin. More work is clearly needed to investigate the potential for phosphorylation in the $MenA_1$-$MenT_1$ system, and what contribution it may make to toxicity or antitoxicity.

In contrast, $MenA_4$ neither detectably modified $MenT_4$ (Figure 5.27), nor physically interacted with the cognate toxin *in vitro* or *in vivo* (Figure 5.24), suggesting that the mechanism of antitoxicity is not conserved among MenA antitoxins. Interestingly, the crystal structure of $MenA_4$ reveals a predicted NTase fold in the CTD, which may indicate an NTP binding site for kinase activity[378]. This is reminiscent of characteristic antitoxin structure; the NTD of $MenA_4$ contains a DNA-binding wHTH fold, important for autoregulation, whilst the CTD contains a motif important for antitoxicity[7,261,378]. However, as mentioned, no $MenT_4$ PTMs were detectable by MS following $MenA_4$ co-expression, suggesting that the predicted NTase fold and putative kinase activity are unrelated.

Considering the lack of detectable $MenA_4$:$MenT_4$ interactions, experimental methodology may have contributed to these observations. It is possible that the use of AEC during protein purification may have caused the disassociation of $MenA_4$:$MenT_4$ complexes formed *in vivo* during co-expressions. Likewise, this may have accounted for the absence of SEC-detectable $MenA_3$:$MenT_3$ interactions from a $MenA_3$+$MenT_3$ co-expression (Figure 5.23). Indeed, AEC had already been used to isolate $MenT_1$ for MS analysis from a $MenA_1$:$MenT_1$ co-expression, where AEC appeared to cause the disassociation of a subset of $MenT_1$ from the $MenA_1$:$MenT_1$ complex (See, *5.7*). Accordingly, our collaborators in the Cai *et al*. study showed through affinity-tagged *in vivo* co-purification experiments in *E. coli* that small but notable fractions of $MenA_1$-$MenT_1$ and $MenA_3$-$MenT_3$ co-purified[98]. In the case of $MenA_1$-$MenT_1$, this is consistent with the crystal structure and SEC data presented here. However, in contrast to $MenA_1$-$MenT_1$, co-incubating $MenA_3$ and $MenA_4$ with their cognate toxins failed to cause *in vitro* complex formation in analytical SEC experiments (Figures 5.23 and 5.24). Considering

the cumulative evidence, a possible conclusion is that $MenA_3$ does not complex with $MenT_3$, but instead interacts transiently to carry out its kinase activity, whereas $MenA_4$ likely does not interact with $MenT_4$ *in vitro* or *in vivo*. Given that $MenA_4$ did not appear to interact with the cognate toxin, nor share $MenA_3$ kinase activity, the question of $MenA_4$ antitoxicity remains unanswered, with a possible non-interacting type IV mechanism still plausible.

The structural characterisation of the MenA-MenT systems presented here showed that this family potentially features a diverse range of TA classifications, including putative type II ($MenA_1$-$MenT_1$), type IV ($MenA_4$-$MenT_4$), and type VII systems ($MenA_3$-$MenT_3$). In contrast, the toxin crystal structures all showed notable similarity, with comparable protein architecture and a shared toxin fold, whilst the characterisation of essential functional residues in $MenT_3$ and $MenT_4$ indicated a conserved active site. However, the cellular target and mechanism of MenT toxicity remained to be explored.

6. MenT toxins target transcription/translation to inhibit bacterial growth

6.1. MenT toxins act distinctly from previously characterised type IV toxins

Due to the original predicted classification of the MenA-MenT TA family as type IV modules[11], the toxin mode of action was initially investigated in the context of previously characterised type IV TA systems. At the time of this study, the two best-characterised type IV systems were AbiEi-AbiEii from *S. agalactiae* and CbtE-CbtA from *E. coli*[81–84]. In the case of the latter, the CbtA toxin inhibits growth by physically interacting with and inhibiting the polymerisation of the FtsZ and MreB cytoskeletal proteins[81,82]. The effect of CbtA toxicity is visibly detectable by microscopy, where toxin expression causes a lemon-shaped *E. coli* cell morphology[82]. Furthermore, the results of earlier toxicity assays suggested that the $MenA_4$-$MenT_4$ system might interact with components important for cell structure; $MenA_4$ expression caused *E. coli* colony morphologies that were larger and more irregular than a vector-only control (Figures 3.2 and 3.5).

Therefore, to explore whether the MenT toxins target cell structure via a similar mechanism of action to CbtA, *E. coli* DH5α were first transformed with plasmids encoding each of the DUF1814 MenT toxins or DUF1814 AbiEii. Expression was induced, and the effects on cell morphology were assessed by laser-scanning confocal microscopy (Figure 6.1). Cell morphologies were generally unremarkable, with most cells broadly comparable to the DH5α pBAD30 vector-only control strain. No lemon-shaped morphologies were detected, with the main morphological discrepancy being a slight enrichment of more filamentous *E. coli* cells following $MenT_3$ expression (Figure 6.1).

These results were unsurprising. AbiEii has been shown to exhibit GTP-specific NTase activity and is homologous to the DUF1814 NTase-like MenT toxins[11,83], whilst Phyre2 searches similarly predicted homology between the MenT toxins and NTase proteins[361]. In contrast, inputting the CbtA AA sequence into the Phyre2 server returned no viable predicted homology hits, with structure modelling consequently particularly poor. Comparing CbtA by pairwise AA sequence alignment to the MenT toxins generates a highest sequence identity of only 27.2%, suggesting little evolutionary overlap with MenT toxins. Indeed, the 124 AA CbtA toxin is considerably smaller than any of the MenT toxins, with a predicted mass of only 13.9 kDa compared to the 21.8 kDa, 23.9 KDa, 31.9 kDa, and 33 kDa for $MenT_1$, $MenT_2$, $MenT_3$ and $MenT_4$, respectively. Furthermore, the variable TA classes present among the MenA-MenT TA modules, for example potentially type II, IV and VII, also support the notion that MenT and CbtA toxins are unrelated, and that MenT toxins likely operate via a distinct mechanism.

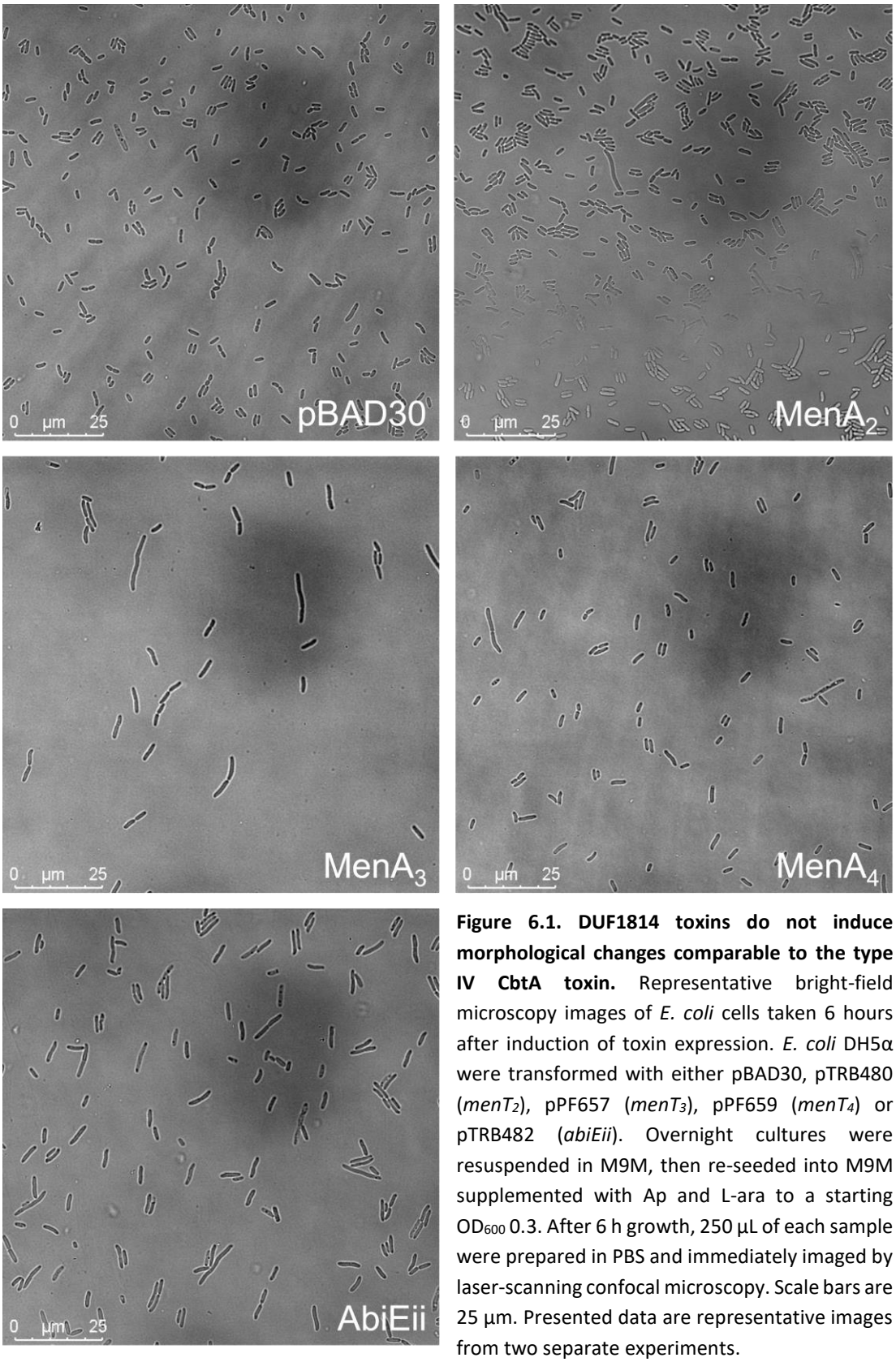


Figure 6.1. DUF1814 toxins do not induce morphological changes comparable to the type IV CbtA toxin. Representative bright-field microscopy images of *E. coli* cells taken 6 hours after induction of toxin expression. *E. coli* DH5α were transformed with either pBAD30, pTRB480 (*menT*$_2$), pPF657 (*menT*$_3$), pPF659 (*menT*$_4$) or pTRB482 (*abiEii*). Overnight cultures were resuspended in M9M, then re-seeded into M9M supplemented with Ap and L-ara to a starting OD_{600} 0.3. After 6 h growth, 250 µL of each sample were prepared in PBS and immediately imaged by laser-scanning confocal microscopy. Scale bars are 25 µm. Presented data are representative images from two separate experiments.

6.2. Structural homologue analyses suggest MenT NTase activity

To gain further insight into the possible toxin mode of action, searches for structural homologues of $MenT_3$ were performed using the DALI server, which compares inputted structures against all entries in the PDB database[360]. Among a number of hits corresponding to NTases, the best match was JHP933 from *Helicobacter pylori* (*H. pylori*) J99, a predicted NTase encoded by the *jhp0933* gene, which is also classified as a DUF1814 protein[404,405].

Pairwise AA sequence alignments between JHP933 and the MenT toxins were performed using EMBOSS Stretcher. This highlighted overall poor sequence identity, with the best scoring alignment between JHP933 and $MenT_3$ at 25.1%. The JHP933 structure (PDB: 4OK0) was next aligned to $MenT_1$, $MenT_3$ and $MenT_4$ (Figure 6.2). JHP933 aligned poorly to both $MenT_1$ (7.06 Å, between 743 atoms) and $MenT_4$ (6.4 Å, between 799 atoms), although several helices and β-strands were roughly overlayed in the aligned structures. Sequence-independent superpositions failed to improve RMSD scores, instead reaffirming poor overall structural similarity ($MenT_1$ = 13.29 Å, between 947 atoms; $MenT_4$ = 14.62 Å, between 1142 atoms). In contrast, JHP933 aligned well to $MenT_3$ with an RMSD of 2.28 Å, between 180 atoms, displaying a high degree of structural similarity, albeit with multiple additional helices resolved in the $MenT_3$ structure (Figure 6.2B).

Despite alignments suggesting structural differences between JHP933 and MenT toxins, the proposed catalytic core overlayed remarkably well, with conserved MenT residues positioned close to their equivalents in JHP933 (Figure 6.2). The secondary structure of JHP933 displays the mixed helix/β-sheet NTD and bundled helix CTD common to the MenT toxins which sandwich the conserved active site, suggesting that the MenT toxins and JHP933 share the same toxin fold. This was affirmed by an analysis of the electrostatic potential of the JHP933 protein surface, which features an electropositive cavity between the NTD and CTD comparable to the MenT toxins (Figure 6.2D).

A follow-up study by one of the groups that originally solved the JHP933 crystal structure sought to characterise its putative NTase activity and pinpoint the JHP933 active site[406]. The authors first performed superpositions of JHP933 with four structures of Cid1, a *Schizoscaccharomyces pombe* uridylyltransferase, complexed with magnesium (Mg) and either ATP, CTP, GTP or UTP[406,407]. From these superpositions, they modelled the location of the respective NTPs to the JHP933 active site and identified key residues predicted to be involved in the NTP-binding interface[406]. Having confirmed that JHP933 did indeed act as an NTase *in vitro*, the authors found that substitution mutants of these residues abolished this

activity[406]. Notably, one of the residues identified as essential to activity was JHP933 D55, which overlays well with the conserved MenT residues D41 ($MenT_1$), D80 ($MenT_3$), and D69 ($MenT_4$), located within the MenT toxin active site (Figure 6.2A to C).

This modelling was repeated by superposing $MenT_3$ with each of the Cid1:NTP structures (Figure 6.3A). These produced particularly poor RMSD scores – 12.906 Å, between 854 atoms (Cid1:ATP), 12.644 Å, between 824 atoms (Cid1:CTP), 12.803 Å, between 830 atoms (Cid1:GTP), and 11.991 Å, between 804 atoms (Cid1:UTP) – which confirmed that $MenT_3$ and Cid1 share no structural similarity. However, by superposing the $MenT_3$ and Cid1:NTP structures, models of $MenT_3$ complexed with ATP, CTP, GTP and UTP were generated (Figure 6.3B). These showed localisation of the NTPs within the $MenT_3$ active site, nestled among the functional residues previously characterised as essential to $MenT_3$ toxicity. Despite the poor RMSD scores of the superposed structures indicating a lack of biological validity, the clustering of these essential residues around the NTP phosphate tails hints at a possible mechanism by which $MenT_3$ would interact with NTPs and carry out its putative NTase activity (Figure 6.3B).

Interestingly, an analysis of the *H. pylori* J99 genome revealed that *jhp0933* lies immediately downstream of gene *jhp0932*, with the coding sequence partially overlapping. The *jhp0932* gene encodes an uncharacterised 156 AA protein, notably smaller than $MenA_2$ (342 AAs), $MenA_3$ (207 AAs), and $MenA_4$ (295 AAs), and larger than $MenA_1$ (68 AAs). Pairwise AA sequence alignments (EMBOSS Stretcher) to the MenA antitoxins showed poor sequence identities, with the best-scoring match between JHP932 and $MenA_3$ at 23.5%. In spite of the low sequence identity, this characteristic presentation of TA genetic architecture, coupled with the NTase-like DUF1814 classification for JHP933, suggests that JHP933 may belong to the MenT toxin family and might function as a TA system *in vivo* in *H. pylori*.

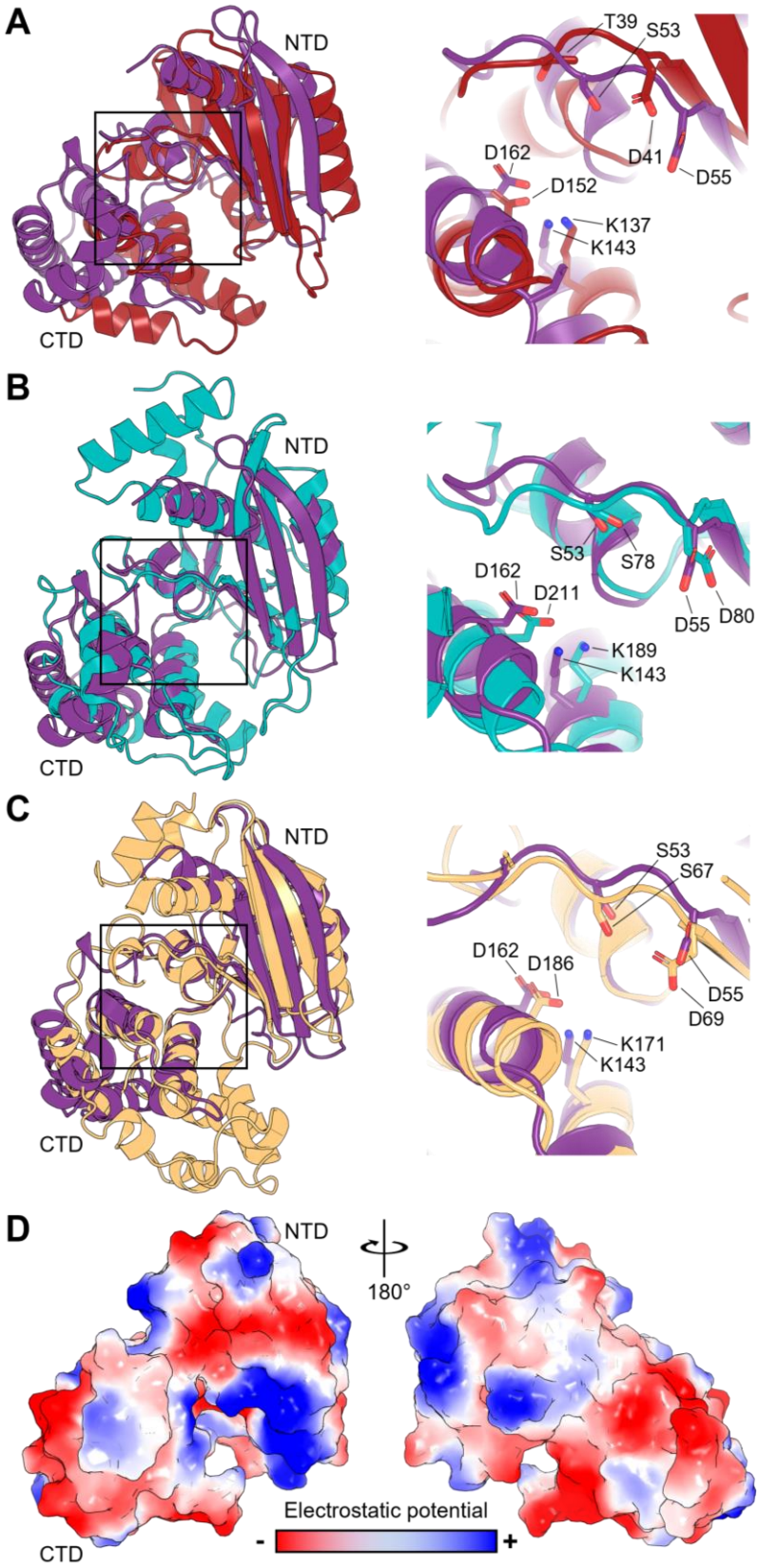


Figure 6.2. MenT$_1$, MenT$_3$, and MenT$_4$ alignments with *H. pylori* JHP933 reveal a conserved active site. Alignment of JHP933 (PDB: 4OK0) as a violet purple cartoon representation to **(A)** MenT$_1$ (protomer A (MenT$_1$ crystal structure), fire brick), **(B)** MenT$_3$ (teal), and **(C)** MenT$_4$ (light orange). **(D)** JHP933 surface electrostatics, presented as front and back views rotated on *y* by 180°; red represents electronegative potential, blue represents electropositive potential.

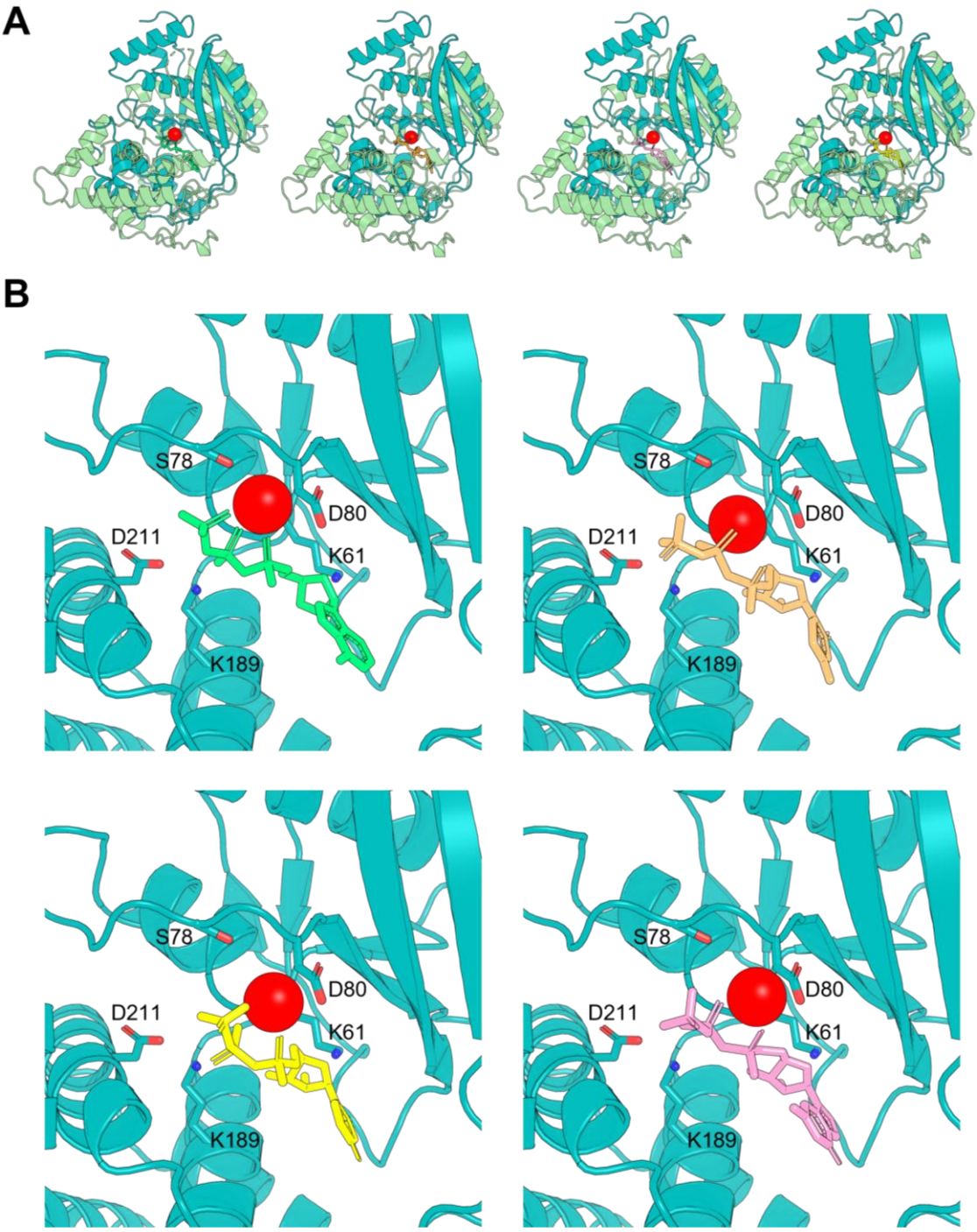


Figure 6.3. MenT$_3$-NTP modelling suggests localisation within the MenT$_3$ active site. (A) MenT$_3$ (teal) structural superpositions with Cid1 (pale green) bound, from left to right, to ATP and Mg (PDB: 4FHY), CTP and Mg (PDB: 4FHV), GTP and Mg (PDB: 4FHW), and UTP and Mg (PDB: 4FH5). **(B)** Close-up views of the MenT$_3$ active site from superpositions in (A), with the Cid1 structure removed. Clockwise from top left: MenT$_3$ and ATP (lime green), MenT$_3$ and CTP (bright orange), MenT$_3$ and GTP (pink), and MenT$_3$ and UTP (yellow). Mg ions are coloured red; structures are displayed as cartoon representations. Key residues essential for MenT$_3$ toxicity are labelled and displayed as sticks.

6.3. $MenT_3$ and $MenT_4$ toxins inhibit protein synthesis *in vitro*

Whilst the functional and structural characterisation of the MenA-MenT TA systems was underway, collaborators from the Fineran group were further investigating the AbiEi-AbiEii TA system from *Serratia* sp. ATCC 39006 (*Serratia*). They communicated, and later published, the results from RNA-seq experiments assessing the transcriptomic response of *Serratia* to AbiEii expression, where the major category of differentially expressed genes following AbiEii expression were tRNA-encoding genes[408]. As a result, attention turned towards a possible role for the DUF1814 MenT toxins in translation inhibition.

To test whether the MenT toxins inhibit protein synthesis, the PURExpress *In Vitro* Protein Synthesis Kit (New England Biolabs) was used. The PURExpress kit is an *in vitro* cell-free expression assay which contains the reconstituted components essential for *E. coli* transcription and translation. The kit allows for protein synthesis from template DNA in a simple one-pot reaction, which can in turn be analysed and quantified by SDS-PAGE. This therefore provided an efficient means to test the effect of MenT toxins on protein synthesis *in vitro*; by adding purified MenT toxins to reactions, the overall amount of protein synthesis could be measured and compared to a no-toxin control. Only $MenT_3$ and $MenT_4$ were tested for inhibitory activity, as no other MenT toxins inhibited growth in *E. coli* (See, *3.4*).

In vitro cell-free expression assays were performed as described in Materials and Methods, *2.13*. The provided dihydrofolate reductase (DHFR)-encoding DNA template was used as an indicator for *in vitro* protein synthesis activity. To begin with, purified $MenT_3$ and $MenT_3$ SEP protein was added to reactions in increasing concentrations and the relative levels of DHFR protein synthesis were analysed by SDS-PAGE (Figure 6.4). Consistent with predictions, the addition of $MenT_3$ inhibited *in vitro* protein synthesis, with the maximum 1 µM toxin concentration sufficient to almost completely abolish activity (Figure 6.4A). Interestingly, the addition of $MenT_3$ SEP protein similarly inhibited protein synthesis, despite the presence of the $MenT_3$ S78 phosphoserine linked to $MenA_3$-mediated antitoxicity (Figure 6.4B). However, inhibition was weaker than that of non-phosphorylated $MenT_3$; densitometry analyses showed that at 0.25 µM toxin concentration, protein synthesis activity in the presence of native $MenT_3$ was reduced approximately 2-fold compared to $MenT_3$ SEP (Figure 6.4C).

The effect of $MenT_4$ on *in vitro* protein synthesis was similarly tested, to determine whether the two MenT toxins that were functional in *E. coli* share a conserved mode of action (Figure 6.5). $MenT_4$ also inhibited expression, though its toxicity appeared less potent that of $MenT_3$; full inhibition of protein synthesis was observed at 5 µM, compared to only 1 µM for $MenT_3$.

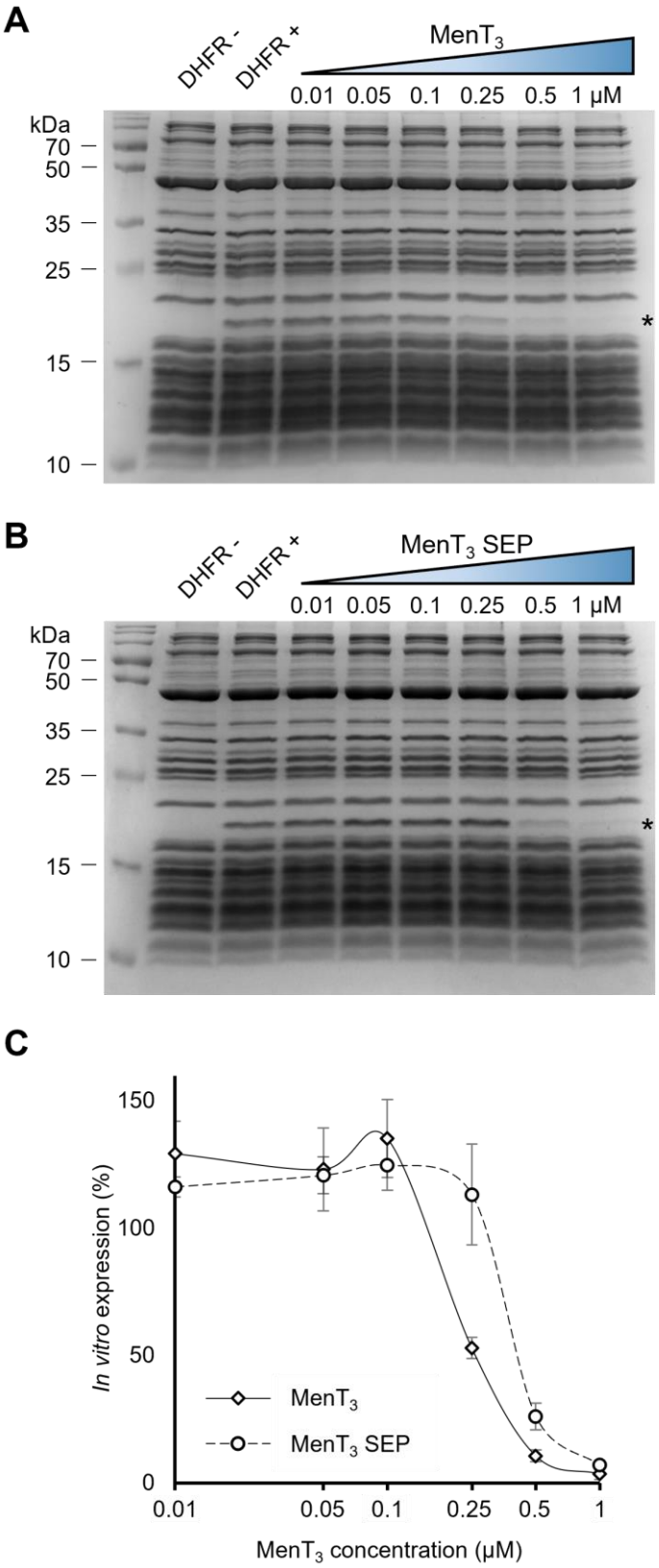


Figure 6.4. MenT$_3$ inhibits protein synthesis in a concentration-dependent manner. *In vitro* cell-free expression assays measuring levels of DHFR protein (black asterisk) produced in the absence or presence of increasing concentrations of **(A)** MenT$_3$ toxin, or **(B)** MenT$_3$ SEP toxin. Samples were separated by SDS-PAGE and stained with InstantBlue. **(C)** Densitometry analysis of cell-free expression assays represented in (A) and (B). Densitometry of DHFR produced in the presence of toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

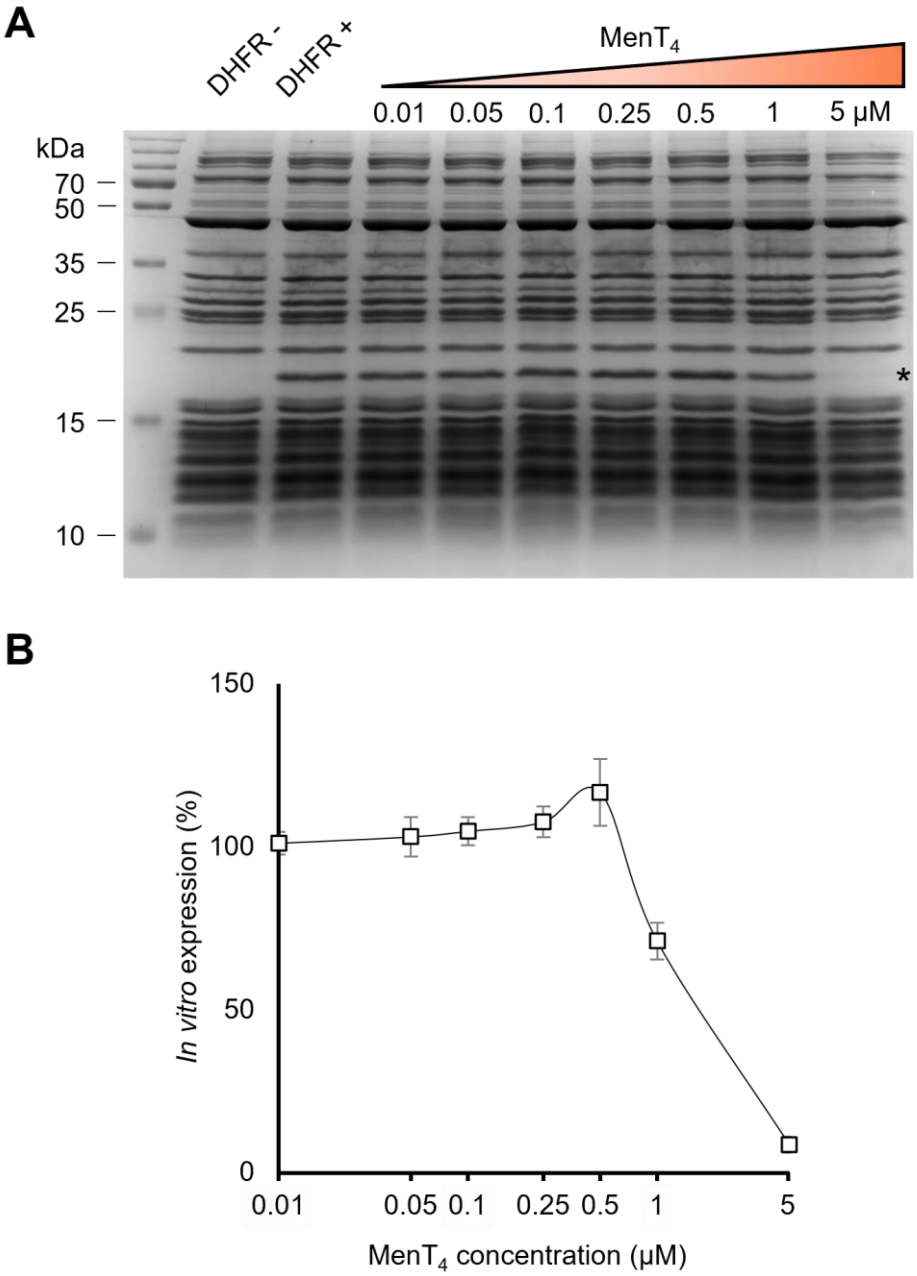


Figure 6.5. MenT$_4$ inhibits protein synthesis in a concentration-dependent manner. (A) *In vitro* cell-free expression assays measuring levels of DHFR protein (black asterisk) produced in the absence or presence of increasing concentrations of MenT$_4$ toxin. Samples were separated by SDS-PAGE and stained with InstantBlue. **(B)** Densitometry analysis of cell-free expression assays represented in (A). Densitometry of DHFR produced in the presence of toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

6.4. $MenT_3$ active site mutants abolish *in vitro* toxicity

Specific residues within the $MenT_3$ active site were next mutated to determine their functional importance *in vitro*. The $MenT_3$ residues previously chosen for alanine substitution and tested by toxicity and antitoxicity assays were carried forward for *in vitro* cell-free expression assays. Expression constructs were first generated encoding each of the $MenT_3$ mutants fused to an N-terminal His_6-SUMO tag. Plasmids pTRB571 ($menT_3^{(Q185A)}$), pTRB572 ($menT_3^{(Q188A)}$), pTRB593 ($menT_3^{(D80A)}$), and pTRB594 $menT_3^{(D211A)}$ were produced by SDM, using pTRB517 (His_6-*SUMO*-$menT_3$) as a template. Attempts to generate $MenT_3^{(K61A)}$ and $MenT_3^{(K189A)}$ expression constructs via SDM were unsuccessful, therefore $menT_3^{(K61A)}$ and $menT_3^{(K189A)}$ were amplified by PCR from pTRB559 and pTRB562, then LIC cloned into pTRB550 to create pTRB576 and pTRB577, respectively. *E. coli* ER2566 were next co-transformed with pPF656 and either pTRB571, pTRB572, pTRB576 or pTRB577, or pRARE and either pTRB593 or pTRB594. $MenT_3$ mutant proteins were expressed and purified as per Materials and Methods *2.6* and *2.7*. *In vitro* cell-free expression assays were then performed assessing the relative amount of DHFR protein synthesis in the absence or presence of 1 μM $MenT_3$ WT or $MenT_3$ mutant protein (Figure 6.6). Consistent with the results from earlier toxicity assays (Figure 5.5A), $MenT_3^{(K61A)}$, $MenT_3^{(D80A)}$, $MenT_3^{(K189A)}$, and $MenT_3^{(D211A)}$ failed to inhibit protein synthesis compared to $MenT_3$ WT protein, highlighting the essentiality of these residues to $MenT_3$ toxicity. In contrast, the Q185A and Q188A substitution mutants retained toxicity, inhibiting DHFR synthesis to a level broadly comparable to $MenT_3$ WT (Figure 6.6).

6.5. $MenA_3$ and $MenA_4$ lack *in vitro* antitoxicity

Co-expression of the $MenA_3$ and $MenA_4$ antitoxins was sufficient to neutralise the toxicity of the cognate $MenT_3$ and $MenT_4$ toxins *in vivo* in *E. coli* (See, *3.4*). Subsequently, the $MenT_3$ and $MenT_4$ toxins were shown to inhibit protein synthesis *in vitro* (Figures 6.4 and 6.5). As a result of these findings, the antitoxicity of $MenA_3$ and $MenA_4$ was next assessed *in vitro* via cell-free expression assays. DHFR protein synthesis reactions were supplemented with either purified toxin protein alone, providing a positive control of inhibition, or supplemented with a mix of toxin and increasing concentrations of the cognate MenA antitoxin. Interestingly, neither of the antitoxins displayed any *in vitro* antitoxicity. Supplementation with the $MenA_3$ antitoxin failed to neutralise the toxicity conferred by either $MenT_3$ WT (Figure 6.7) or $MenT_3$ SEP (Figure 6.8), even at an antitoxin concentration five times that of the toxin, whilst $MenA_4$ was similarly unable to restore protein synthesis at any concentration tested (Figure 6.9).

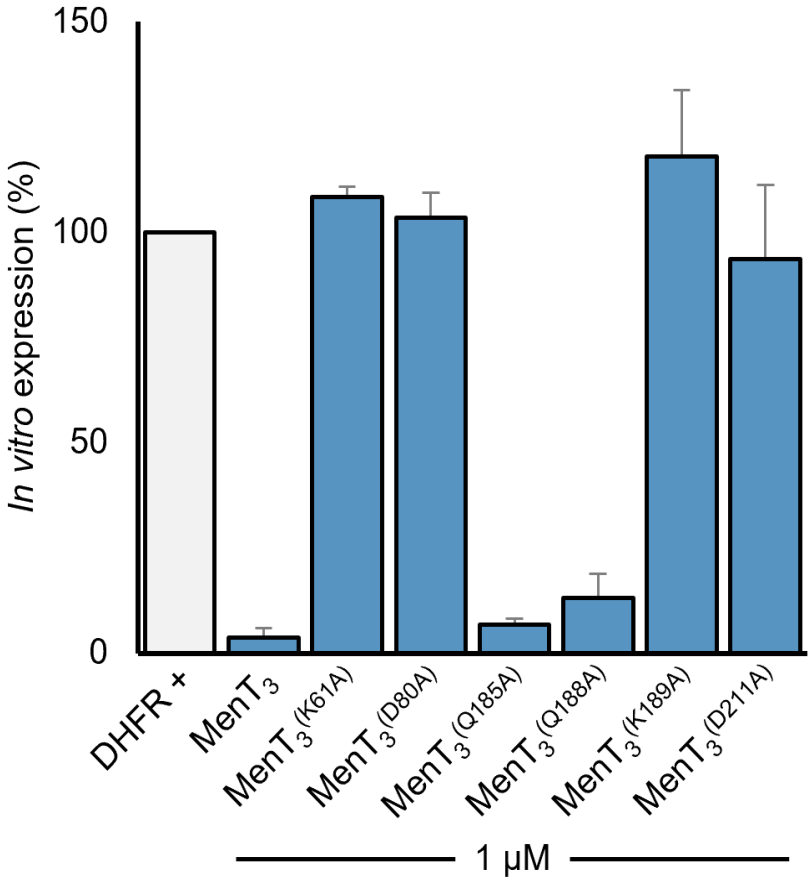


Figure 6.6. Conserved MenT$_3$ active site substitution mutants restore protein synthesis. Densitometry analysis of *in vitro* cell-free expression assays measuring levels of DHFR protein produced in the absence or presence of 1 μM MenT$_3$ WT or mutant proteins. Sample reactions were first separated by SDS-PAGE, prior to densitometry analysis. Densitometry of DHFR produced in the presence of toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

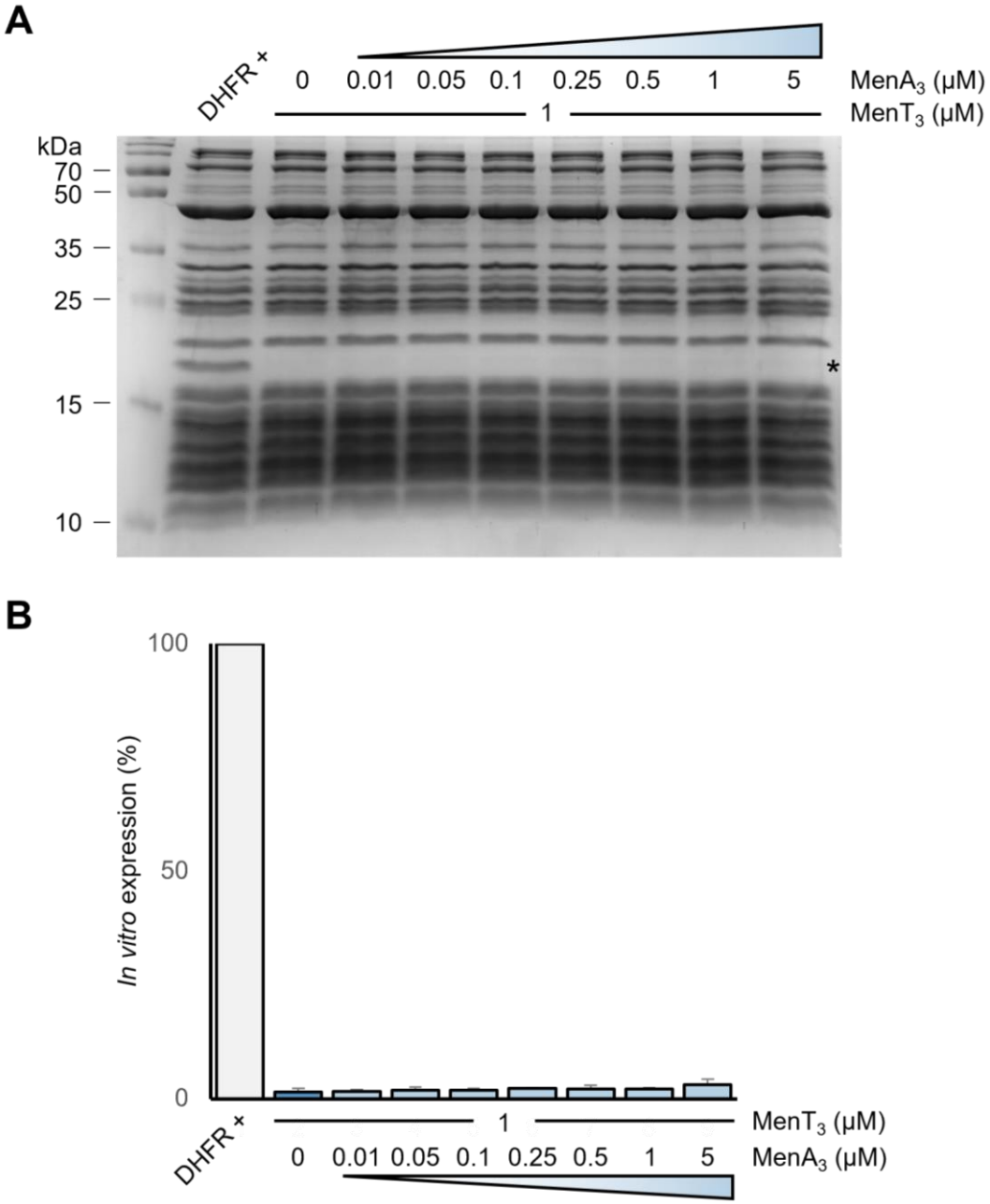


Figure 6.7. MenA$_3$ fails to prevent MenT$_3$ toxicity *in vitro*. (A) *In vitro* cell-free expression assay measuring levels of DHFR protein (black asterisk) produced in the absence or presence of 1 μM MenT$_3$ toxin, and in the absence or presence of increasing concentrations of MenA$_3$ antitoxin. Samples were separated by SDS–PAGE and stained with InstantBlue. **(B)** Densitometry analysis of cell-free expression assays represented in (A). Densitometry of DHFR produced in the presence of antitoxin and/or toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

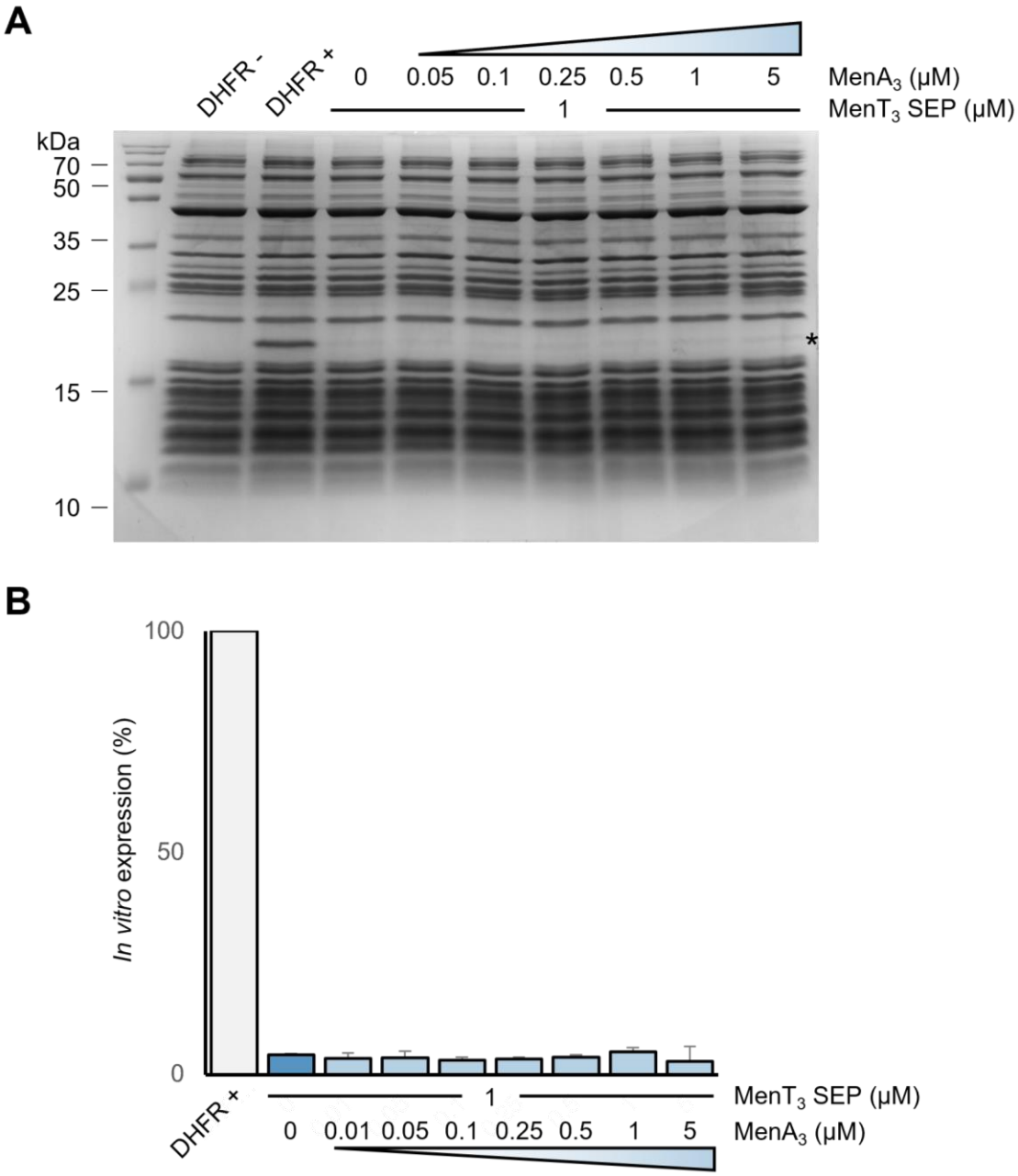


Figure 6.8. $MenA_3$ fails to prevent $MenT_3$ SEP toxicity *in vitro*. (A) *In vitro* cell-free expression assay measuring levels of DHFR protein (black asterisk) produced in the absence or presence of 1 μM $MenT_3$ SEP toxin, and in the absence or presence of increasing concentrations of $MenA_3$ antitoxin. Samples were separated by SDS–PAGE and stained with InstantBlue. **(B)** Densitometry analysis of cell-free expression assays represented in (A). Densitometry of DHFR produced in the presence of antitoxin and/or toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

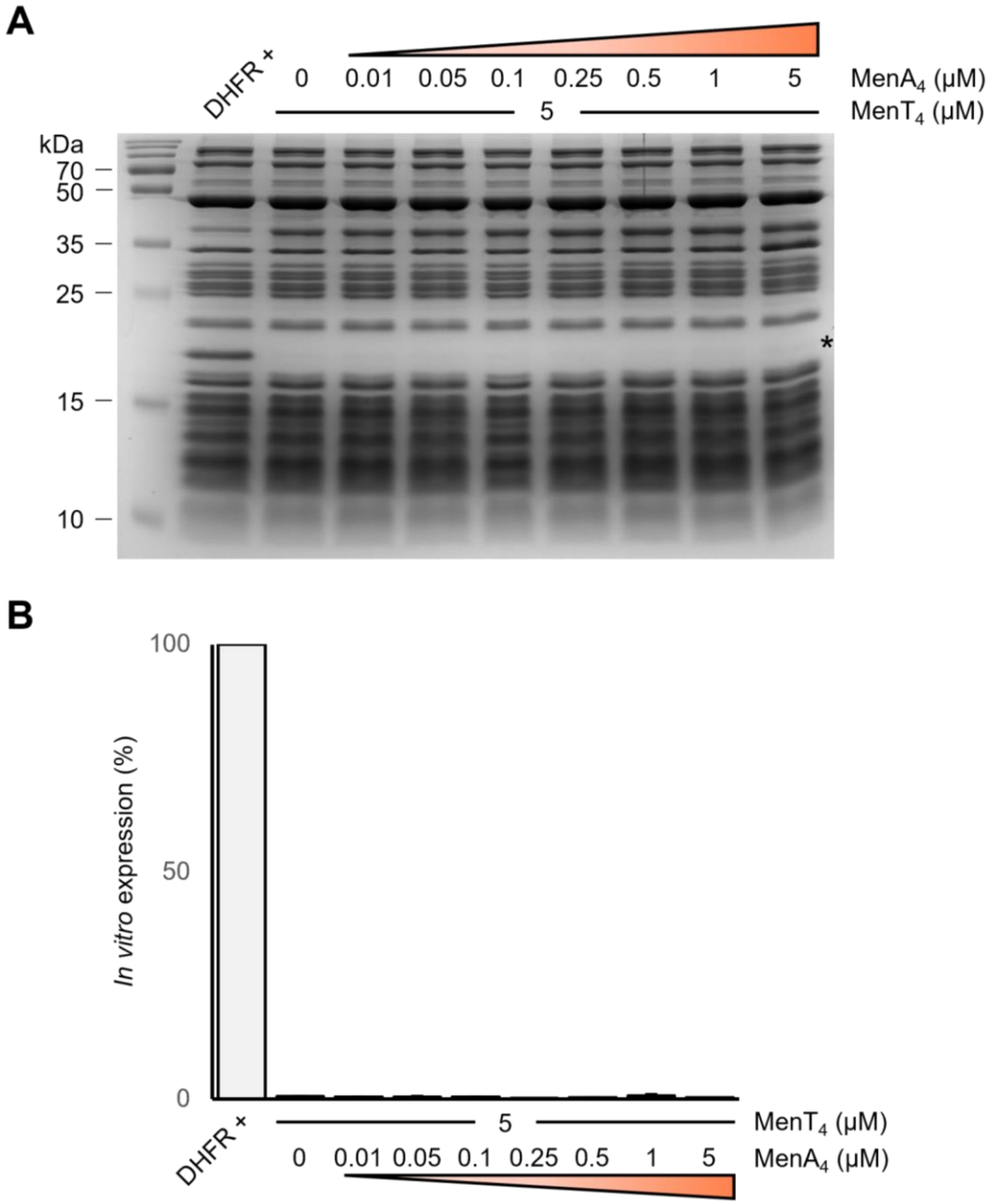


Figure 6.9. MenA$_4$ fails to prevent MenT$_4$ toxicity *in vitro*. (A) *In vitro* cell-free expression assay measuring levels of DHFR protein (black asterisk) produced in the absence or presence of 5 μM MenT$_4$ toxin, and in the absence or presence of increasing concentrations of MenA$_4$ antitoxin. Samples were separated by SDS–PAGE and stained with InstantBlue. **(B)** Densitometry analysis of cell-free expression assays represented in (A). Densitometry of DHFR produced in the presence of antitoxin and/or toxin was measured based on the relative SDS-PAGE band density compared to DHFR produced alone, which was standardised as 100% *in vitro* expression. Plotted data represent the mean ± standard deviation (≥ 3 replicates).

6.6. Discussion

The cellular processes targeted by the homologous $MenT_3$ and $MenT_4$ toxins were investigated. Structural analyses suggested potential roles for the MenT toxins as NTases targeting electronegative substrates, such as nucleotides or tRNAs. Subsequent *in vitro* expression assays demonstrated that the activity of both $MenT_3$ and $MenT_4$ impaired protein synthesis by targeting a component or process within the transcription/translation pathway. The assay used allowed for an efficient visual and densitometric analysis of whether toxin activity inhibited protein production. However, the method used – adding a DNA template to a mix of purified *E. coli* transcription and translation components – did not allow for the discrimination between which process – transcription or translation – was targeted. For future experiments, omitting the DNA template and adding mRNA template directly to the assay could bypass transcription and help to determine whether translation is specifically targeted by MenT toxins.

Transcription and translation are not novel targets of TA systems. The type VII *E. coli* Hha toxin represses the transcription of regulatory rare tRNAs, which in turn inhibits the production of type I fimbriae[95]. Separately, the Ccdb toxin from the *E. coli* F plasmid, when complexed with GyrA and bound to DNA, can block the passage of RNA polymerase to disrupt transcription[110]. Meanwhile, many toxins have been shown to interfere with translation, and the mechanisms of toxicity are highly variable: some toxins are RNases that target mRNAs for degradation, such as *E. coli* MazF which specifically cleaves free mRNAs at the A▼CA sequence[71]. Other mRNase toxins such as *E. coli* RelE and YoeB interfere with mRNA-ribosome processing[123,124], whereas *E. coli* RatA blocks 70S ribosome formation[125]. The work presented here adds to this expanding library of transcription/translation inhibitors, clearly demonstrating that both $MenT_3$ and $MenT_4$ inhibit expression *in vitro*. However, the specific toxin target and mechanism of toxicity could not be elucidated by this work alone.

Fortunately, the results presented here also contributed to a broader collaborative effort characterising MenT toxins, which culminated in the 2020 report by Cai *et al*.[98]. In this study, we elaborated on the mechanistic detail underpinning $MenT_3$ activity, confirming the prediction from structural analyses that $MenT_3$ inhibits translation via tRNA NTase activity. More specifically, we first showed that $MenT_3$ blocked tRNA charging *in vivo*, before subsequent *in vitro* work elucidated the molecular mechanisms at play. We demonstrated that $MenT_3$ adds pyrimidines to the 3'-CCA acceptor stem of uncharged tRNAs *in vitro*, whilst showing strong substrate specificity for the 3'-CCA motif of $tRNA^{Ser}$ (Figure 6.10). This activity

prevents tRNA charging with cognate AAs, thereby disrupting translation, and this is proposed to account for the growth inhibition seen *in vivo* in *E. coli*, *M. smegmatis* and *M. tuberculosis*[98].

The range of toxins that target tRNAs is extensive. In *M. tuberculosis*, the type II VapC toxins are RNases which cleave initiator tRNAs to inhibit translation[130]. Separately, AtaT from *E. coli* and TacT from *S*. Typhimurium function as tRNA acetyltransferases which interfere with the activity of already charged tRNAs; AtaT specifically acetylates the amine group of initiator $tRNA^{fMet}$, which prevents formation of the translation initiation complex[70], whilst TacT acetylates the primary amine group of charged tRNAs, blocking translation processivity and preventing protein synthesis[132]. Furthermore, $MenT_3$ is not the only toxin that modifies the tRNA 3'-CCA motif. The FaRel2, PhRel, PhRel2 and CapRel toxins, members of the toxSAS RSH family (small alarmone synthetase RelA-SpoT homologue toxins), have been found to inhibit protein synthesis by pyrophosphorylating the 3'-CCA of uncharged tRNAs, preventing their aminoacylation[409]. In addition, the type VII HEPN RNase toxin chromosomally encoded by the cyanobacterium *Aphanizomenon flos-aquae* cleaves the 3'-CCA motif and fourth discriminator nucleotide of various tRNAs to prevent tRNA charging[410]. Clearly, translation in general is a popular target of TA systems, due presumably to the abundance of components involved in the process, and its essentiality to cell viability. The fact that even within this vast process, the same component can be targeted by multiple toxins via distinct mechanisms perfectly illustrates the sheer diversity among TA systems, and might reflect that each system is also activated by distinct conditions.

In light of the results presented here, and those reported by Cai *et al*., 2020[98], $MenT_3$ provides yet another example of the diverse mechanisms by which TA systems can target translation. It remains to be seen whether the $MenT_3$ mechanism is conserved among DUF1814 MenT toxins. While $MenT_1$ and $MenT_4$ share structural similarity to $MenT_3$, and $MenT_4$ was here shown to inhibit protein synthesis *in vitro*, the molecular mechanism responsible for $MenT_4$ toxicity has not yet been explored. The fact that $MenT_4$ was less toxic compared to $MenT_3$ *in vitro* might imply mechanistic differences; potentially reflecting the reduced availability of a distinct target or substrate.

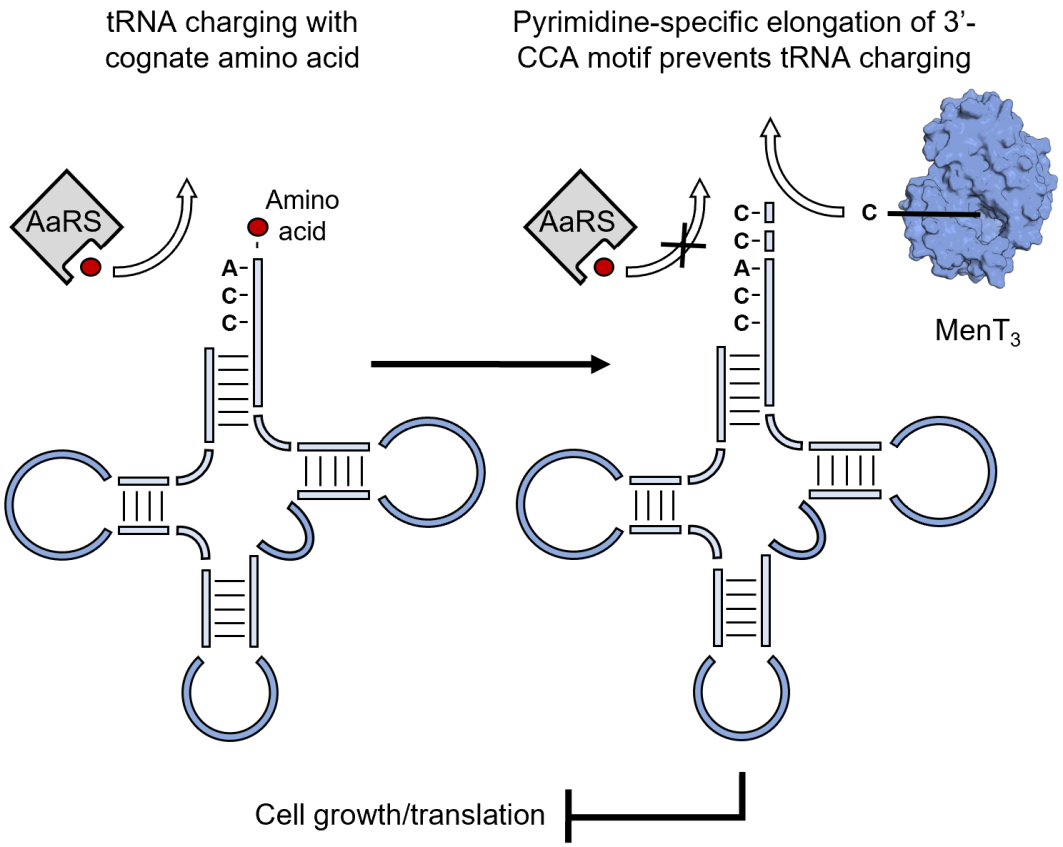


Figure 6.10. Schematic diagram of the MenT$_3$ toxin mechanism of action. MenT$_3$ elongates the 3'-CCA motif of specific tRNAs via pyrimidine transfer, preventing their charging with cognate AAs by aminoacyl-tRNA synthetases (AaRSs). This in turn interferes with translation and inhibits bacterial cell growth.

Despite indicating that translation is the likely cellular process targeted by MenT$_3$ and MenT$_4$, the results from the cell-free expression assays described above raised further unanswered questions. Whilst purified native MenT$_3$ inhibited protein synthesis, the addition of phosphorylated MenT$_3$ SEP also affected *in vitro* expression. This was notable due to the biological context of the MenT$_3$ SEP protein. The phosphorylated serine at S78, present in MenT$_3$ SEP protein produced alongside MenA$_3$ co-expression, has herein been proposed to be relevant to the antitoxicity mechanism. Moreover, and as mentioned earlier, a separate study by Yu *et al.*, 2020, reported that MenA$_3$ acts as a type VII antitoxin, where it phosphorylates MenT$_3$ at residue S78 to neutralise toxicity[99]. The fact that phosphorylated MenT$_3$ SEP still inhibited *in vitro* protein synthesis is therefore striking, not least because MenT$_3$ SEP toxicity is only approximately 0.5-fold weaker than non-phosphorylated native MenT$_3$.

This raises questions as to the actual antitoxic effect of $MenT_3$ S78 phosphorylation and what physiological role this PTM performs. The study by Yu *et al*. reported that phosphorylated $MenT_3$ possessed weaker binding affinity in NTP-binding assays compared to non-phosphorylated $MenT_3$, which was in turn correlated to diminished toxicity[99]. However, this conclusion lacked rigorous testing. The authors' NTP binding assay in fact revealed that phosphorylated $MenT_3$ still retained binding affinity that was only approximately 0.5-fold weaker than non-phosphorylated $MenT_3$[99]. This was comparable to the difference in toxicity between $MenT_3$ SEP and native $MenT_3$ in the *in vitro* cell-free expression assays presented here (Figure 6.4), indicating that activity is still clearly retained despite phosphorylation. Furthermore, they concluded that phosphorylated $MenT_3$ was non-toxic based on an imperfect toxicity assay in *E. coli*, where they relied on a $MenT_3$ S78D mutant to mimic S78 phosphorylation[99].

In comparison, the *in vitro* cell-free expression assays described above tested purified phosphorylated $MenT_3$ SEP protein produced from a $MenA_3$+$MenT_3$ co-expression (Figure 6.4). These results showing inhibition of protein synthesis by $MenT_3$ SEP, allied to the $MenT_3$-NTP binding activity reported by Yu *et al*.[99], suggest that $MenT_3$ S78 phosphorylation is not sufficient alone to completely abrogate $MenT_3$ activity. Perhaps S78 phosphorylation serves merely to diminish $MenT_3$ toxicity, but not entirely abolish it, by disrupting NTP binding at the active site. This could provide a basal level of toxicity regardless of $MenA_3$ expression levels which might favour *M*. *tuberculosis* adaptation under sudden environmental changes. This, however, seems unlikely, given the strong antitoxicity demonstrated by $MenA_3$ in *E. coli*, where even uninduced leaky expression of $MenA_3$ sufficiently neutralised $MenT_3$ toxicity (Figures 3.7 and 3.8). It seems more probable that a separate cofactor is involved in $MenA_3$ antitoxicity which acts on the presence of the $MenT_3$ phosphoserine. For example, in Gram-positive bacteria, the McsB kinase tags proteins for Clp protease degradation by phosphorylating arginine residues[396]. Separately, phosphorylation of the anti-σ factor RseA in *M*. *tuberculosis* marks it for degradation by the ClpC1P2 protease, which frees the cognate alternative σ factor SigE and thereby activates the SigE regulon[411]. Whilst purely speculative, a similar mechanism may be involved in $MenA_3$ antitoxicity *in vivo*, where $MenT_3$ is tagged for proteolysis by $MenA_3$-mediated phosphorylation.

Alternatively, the retained toxicity of $MenT_3$ SEP protein, albeit at a lower level than that of native $MenT_3$, might simply reflect the presence of a small proportion of non-phosphorylated $MenT_3$ protein in the sample. Indeed, this was indicated by MS analyses, which showed that $MenT_3$ protein purified following $MenA_3$ co-expression contained two distinct species: one

species, the dominant analyte, was phosphorylated $MenT_3$; the second species, making up a smaller proportion of the overall sample, represented non-phosphorylated $MenT_3$ (Figure 5.26B). The reasons for this heterogeneous $MenT_3$ protein sample are unclear. If the $MenA_3$ antitoxin is more labile than its cognate toxin (not demonstrated by this body of work), as is commonly described for canonical TA systems[7], then perhaps $MenT_3$ SEP heterogeneity reflects the speedier degradation of $MenA_3$ protein, and consequently the loss of kinase activity. This would detrimentally affect the ability of $MenA_3$ to match $MenT_3$ turnover and inactivate the toxin when both are ectopically expressed from separate plasmids. The overnight expression methodology used to produce $MenT_3$ SEP protein might also be responsible for sample heterogeneity, by creating a stressful environment characterised by bacteria over-crowding, oxygen and nutrient depletion, and culture acidification, all of which could promote the loss of $MenA_3$.

An alternative biological rationale for the heterogeneous $MenT_3$ SEP population and retained $MenT_3$ SEP toxicity is the activity of innate bacterial phosphatases. Due to the relative stability of Ser/Thr phosphorylation, phosphatase activity is typically required to enzymatically reverse this PTM[412]. In *M*. *tuberculosis*, the only Ser/Thr phosphatase thus far identified is PstP[412]. PstP functions as a global phosphatase with a diverse phosphoproteome, and is implicated in several important physiological roles; it is required for regulated cell division and cell growth, and for the survival of *M*. *tuberculosis* in a murine host[413,414]. PstP might therefore play a role in regulating the $MenA_3$-$MenT_3$ system *in vivo* in *M*. *tuberculosis*, by dephosphorylating $MenT_3$ either spontaneously or condition-dependently. Supporting this hypothesis, Yu *et al*. reported that PstP does in fact dephosphorylate $MenT_3$ *in vitro*, though at a slower rate and with much-reduced activity relative to dephosphorylation of its cognate PknB kinase substrate[99]. They also confirmed that neither $MenA_3$ or $MenT_3$ were capable of dephosphorylating $MenT_3$ *in vitro*, highlighting PstP as the likeliest candidate responsible[99]. In *E. coli*, $MenT_3$ SEP heterogeneity might similarly be achieved during $MenA_3$+$MenT_3$ co-expressions by the action of Ser/Thr phosphatases, for example PphC, PrpA or PrpB[415,416].

The observations regarding $MenT_3$ SEP toxicity relate to another surprising aspect of these assays; namely the lack of antitoxicity exhibited by $MenA_3$ and $MenA_4$ *in vitro*. $MenA_3$ antitoxicity has been demonstrated *in vivo* in *E. coli* (Figures 3.7 and 3.8), *M. smegmatis* and *M*. *tuberculosis*[98]. Furthermore, the Cai *et al*. study also found that co-incubation of $MenA_3$ with $MenT_3$ restricted the NTase activity of $MenT_3$ *in vitro* in tRNA nucleotide transfer assays, but was unable to reverse this activity[98]. This suggests that $MenA_3$ directly inhibits $MenT_3$, rather than enzymatically reversing $MenT_3$ effects, consistent with its reported kinase

activity[99]. These findings made the absence of $MenA_3$ antitoxicity observed in Figure 6.7 all the more intriguing, given the comparable co-incubation methodology involved in both *in vitro* assays. For the expression assays described above (Figure 6.7), purified $MenA_3$ and $MenT_3$ proteins were first mixed together prior to adding to reactions. Similarly, for the *in vitro* tRNA nucleotide transfer assays described by Cai *et al*., 2020, $MenA_3$ and $MenT_3$ were jointly added to the reaction alongside $tRNA^{Ser2}$ and CTP[98].

It is possible that the length of pre-incubation was insufficient for $MenA_3$ activity to take place during *in vitro* expression assays, though given the similar methodology described for both experiments this seems unlikely. As mentioned above, $MenA_3$ might require a secondary cofactor or ligand to mediate antitoxicity, which is absent from the *in vitro* expression assay but present *in vivo* as well as *in vitro* in tRNA nucleotide transfer assays. However, one of the issues that prevents effective troubleshooting of the PURExpress *in vitro* expression assay is the fact that its specific contents are proprietary and unknown. Therefore, directly comparing it against other *in vitro* assays to identify a common denominator is not feasible. Given the results observed elsewhere, however, it is possible that experimental error is instead responsible for the absence of antitoxicity in *in vitro* expression assays. Given that the assays were proven to be functional in the presence of both $MenT_3$ and $MenT_4$ toxins, considering the effect of the toxins on protein synthesis, the common denominator would seem to be the $MenA_3$ antitoxin protein used. To troubleshoot, these assays should be repeated testing antitoxin protein alone as a control, to ensure no adverse effects on protein synthesis due to the addition of $MenA_3$. In addition, despite testing a $MenA_3$ titration against $MenT_3$ toxicity, the ratio of antitoxin to toxin may simply have been insufficient to neutralise $MenT_3$ *in vitro* and not reflective of the *in vivo* ratio. Indeed, several reports have highlighted that antitoxin translation is more efficient than toxin translation, leading to an intracellular antitoxin surplus[14,198,276,303]. As such, repeating experiments using an expanded concentration range of freshly prepared $MenA_3$ protein, as well as testing a range of different duration $MenA_3$ and $MenT_3$ pre-reaction co-incubations, would be an important next step in measuring antitoxicity.

In contrast to $MenA_3$, the mechanism by which $MenA_4$ neutralises $MenT_4$ toxicity is still unknown. However, its antitoxicity has been demonstrated *in vivo* in both *E. coli* (Figure 3.7) and in *M. smegmatis*[98]. Once again, the ratio of antitoxin to toxin tested *in vitro* in expression assays may not have been sufficient to neutralise $MenT_4$ toxicity, and may not be reflective of *in vivo* antitoxin levels. On the other hand, given that $MenA_4$ appears not to add PTMs to $MenT_4$ in a manner comparable to the $MenA_3$-$MenT_3$ antitoxicity mechanism (Figure 5.27),

and does not physically interact with $MenT_4$ as occurs in the $MenA_1$-$MenT_1$ system (Figure 5.24), it is feasible that $MenA_4$-$MenT_4$ could yet act as a type IV TA module. If this were the case, the observation that $MenT_4$ inhibited expression *in vitro*, but $MenA_4$ could not rescue protein synthesis in the same assay, is surprising. $MenA_4$ might therefore act via a non-interacting type IV mechanism to antagonise $MenT_4$ toxicity, but in doing so may also atypically interact with a separate target to that of the cognate toxin. Clearly, further work to explore the $MenT_4$ mechanism and $MenA_4$ antitoxicity is required.

The results described here combined structural and biochemical analyses to show that $MenT_3$ and $MenT_4$ inhibit protein synthesis *in vitro*, likely via an NTase role targeting electronegative tRNAs. Collaborative efforts confirmed these predictions, finding that $MenT_3$ functions as an NTase to transfer pyrimidines to the 3'-CCA motif of uncharged tRNAs[98]. Taken altogether, these results reveal a novel mechanism by which a bacterial toxin targets translation, further diversifying the range of mechanisms by which TA systems disrupt protein synthesis to inhibit cell growth.

7. Final Discussion

7.1. Project overview

The work presented here originally aimed to characterise three putative type IV TA systems from *M*. *tuberculosis*: Rv0837c-Rv0836c, Rv1044-Rv1045, and Rv2827c-Rv2826c. These systems had originally been identified based on homology to the type IV AbiEi-AbiEii system from *S. agalactiae*[11,83]. The Rv0836c, Rv1045, and Rv2826c toxins were classified as NTase-like DUF1814 proteins alongside AbiEii, whereas the cognate antitoxins also shared structural homology with AbiEi[11,83]. This hinted at the presence of a conserved but uncharacterised TA family in *M*. *tuberculosis*. A fourth *M*. *tuberculosis* TA system was later identified which also encoded a DUF1814 toxin, and which was shown to exhibit TA activity in *M. smegmatis*[386]. Based on the work reported here and alongside collaborators, these systems were renamed "mycobacterial AbiE-like NTase antitoxins" (MenA) and "toxins" (MenT) and numbered according to their order in the *M*. *tuberculosis* H37Rv genome: $MenA_1$-$MenT_1$ (Rv0078B-Rv0078A), $MenA_2$-$MenT_2$ (Rv0837c-Rv0836c), $MenA_3$-$MenT_3$ (Rv1044-Rv1045), and $MenA_4$-$MenT_4$ (Rv2827c-Rv2826c).

7.2. Summary of results

This thesis describes the experimental efforts undertaken to characterise the four MenA-MenT TA systems. Phenotypic growth assays showed that $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ function as TA systems in *E. coli*; $MenT_3$ and $MenT_4$ expression caused a dramatic reduction in bacterial growth, whereas co-expression of the cognate antitoxins restored growth. In contrast, $MenA_2$-$MenT_2$ exhibited no characteristic TA activity in *E. coli*, either suggesting its misidentification as a TA system, or possibly the absence of a conserved target. Later tests performed by collaborators in the Cai *et al*. study established that $MenT_2$ also failed to inhibit growth in *M. smegmatis*[98], however its activity in *M*. *tuberculosis* remains untested.

The mechanism of MenT toxicity was explored via structural and biochemical characterisation. These efforts initially led to the determination of the $MenT_1$, $MenT_3$, and $MenT_4$ crystal structures, which revealed that the MenT toxins share a conserved toxin fold and active site. Conserved $MenT_3$ and $MenT_4$ active site residues were identified and functionally characterised using mutagenesis studies, which highlighted their essentiality for toxin activity. Candidate residues for toxin activity were also identified in $MenT_1$; currently these remain uncharacterised but are predicted to be similarly vital to $MenT_1$ toxicity. The $MenT_3$ and $MenT_4$ toxins were further characterised biochemically using *in vitro* cell free expression assays, where both $MenT_3$ and $MenT_4$ were shown to inhibit protein synthesis.

Based on homology searches, MenT toxicity was putatively linked to NTase activity, potentially targeting tRNAs; this was separately confirmed by collaborators[98].

Lastly, a multi-discipline approach was used to characterise the diverse TA types that comprise the MenA-MenT family. X-ray crystallography was used to elucidate the complexed $MenA_1$:$MenT_1$ crystal structure, which showed that $MenA_1$ binds two $MenT_1$ protomers asymmetrically, which is predicted to block substrate access to the toxin active site. $MenA_1$:$MenT_1$ complex formation was confirmed by SEC protein interaction studies, highlighting a putative type II mechanism (Figure 7.1A).

Separately, a second crystal structure of $MenT_3$ contained a phosphoserine at residue S78 which was linked to the co-expression of $MenA_3$. SEC showed no detectable interactions between $MenA_3$ and $MenT_3$, whilst MS analyses confirmed that phosphorylation was mediated by the antitoxin, suggesting that $MenA_3$ functions as a kinase to neutralise $MenT_3$. This was independently confirmed by a different study[99], highlighting a novel mechanism of antitoxicity and placing $MenA_3$-$MenT_3$ within the exclusive type VII TA classification (Figure 7.1B).

In contrast, the expression of $MenA_4$ alongside $MenT_4$ caused no demonstrable toxin modifications, such as phosphorylation, comparable to $MenA_3$-$MenT_3$. Promoter activity assays *in vivo* in *E. coli* established that $MenA_4$ negatively autoregulates transcription from the cognate promoter, whilst SEC showed no interactivity between $MenA_4$ and $MenT_4$. Taken together, these results suggest that a non-interacting type IV mechanism, consistent with AbiEi-AbiEii and aligned with the original prediction for these systems, is still possible (Figure 7.1C).

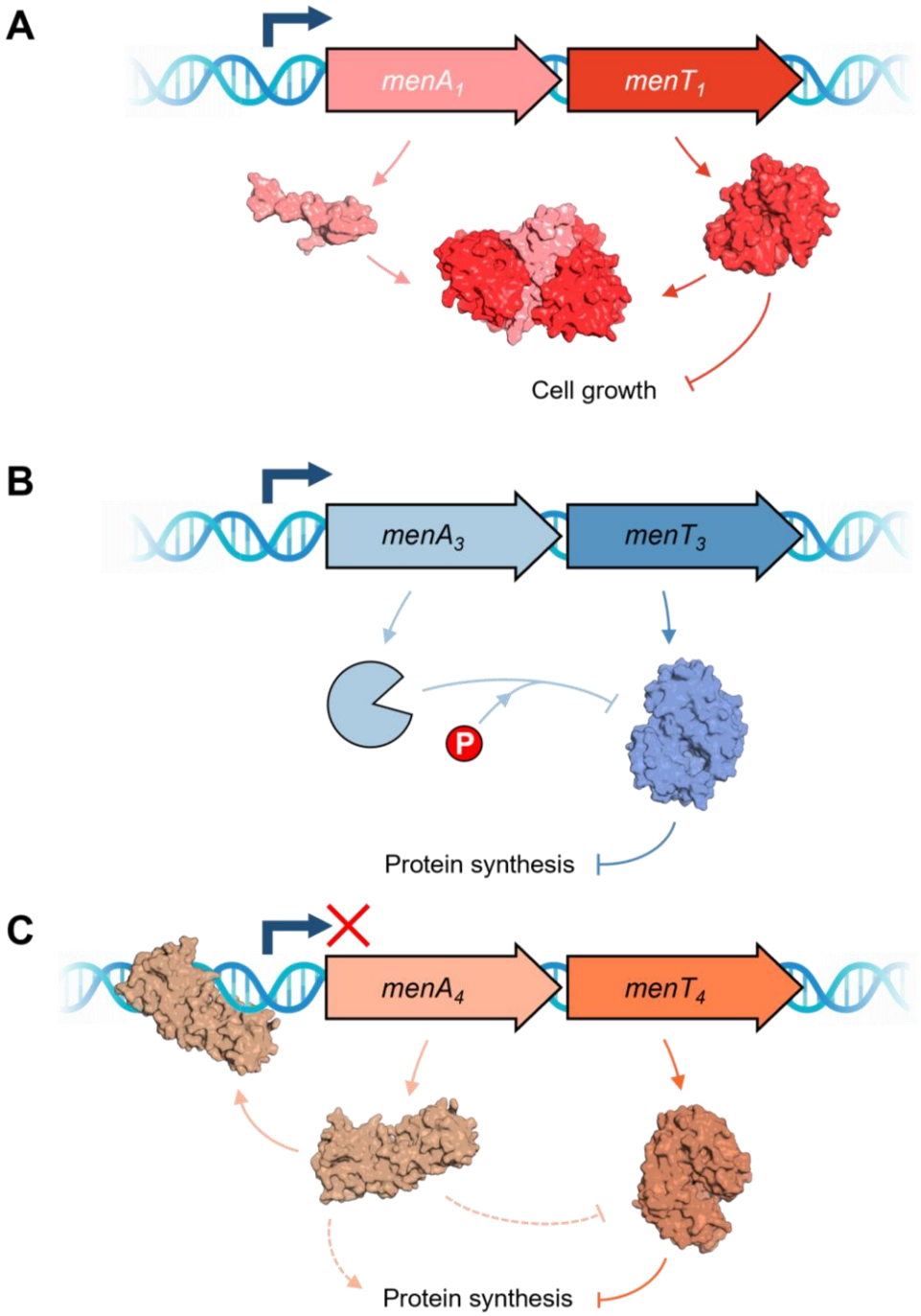


Figure 7.1. Schematic summary of results: diverse classes comprise the MenA-MenT family. (A) $MenT_1$ inhibits cell growth through an unknown mechanism. Antitoxicity occurs via a putative type II mechanism; $MenA_1$ binds two $MenT_1$ protomers, forming an inhibitory $MenA_1$:$MenT_1$:$MenA_1$ complex. **(B)** $MenT_3$ inhibits protein synthesis by preventing the aminoacylation of tRNAs[99]. The $MenA_3$ antitoxin acts through a type VII mechanism, phosphorylating $MenT_3$ S78 to neutralise toxicity. **(C)** $MenT_4$ inhibits protein synthesis through an unknown mechanism. $MenA_4$ (PDB: 1ZEL) neutralises $MenT_4$ toxicity through an unknown non-interacting mechanism. $MenA_4$ also negatively autoregulates the cognate promoter to repress transcription of the *menA₄-menT₄* operon, suggesting a potential type IV mechanism. Solid lines indicate known activity; dashed lines represent unknown activity.

7.3. Model of MenT$_3$ toxicity

Our collaborators in France showed that MenT$_3$ is also toxic in *M. smegmatis* and *M. tuberculosis*, and that the MenA$_3$ antitoxin is essential for *M. tuberculosis* growth, which was reported in the Cai *et al*. paper published in 2020[98]. This toxicity was attributed to the NTase activity of MenT$_3$, which confirmed the structural predictions of MenT NTase activity discussed earlier in this thesis. MenT$_3$ transferred pyrimidines to the 3'-CCA motif of tRNAs *in vitro*[98], preferentially targeting *M. tuberculosis* tRNA^{Ser} isoacceptors, though also observed to modify *M. tuberculosis* tRNA^{Leu2} (Figure 7.2).

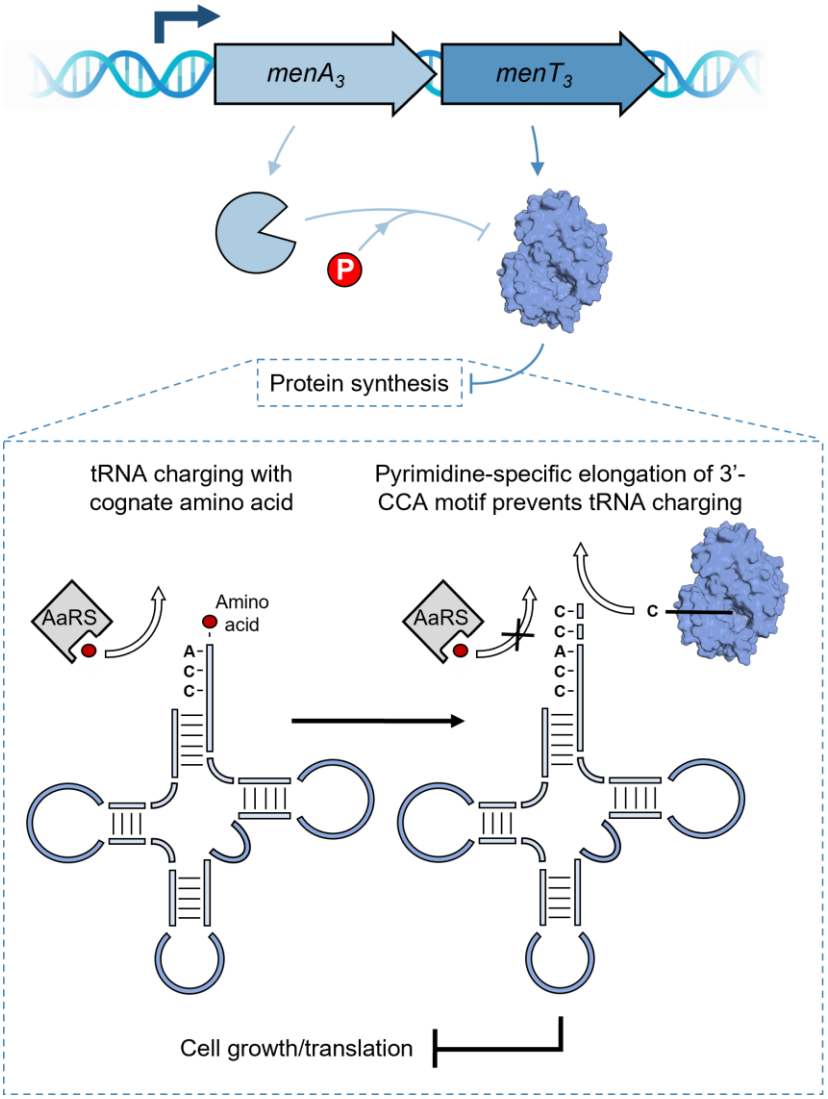


Figure 7.2. Schematic model of MenT$_3$ activity. MenT$_3$ adds pyrimidines to the 3'-CCA motif of specific tRNAs, preventing aminoacylation and inhibiting translation; this is proposed to cause the growth inhibition resulting from MenT$_3$ expression. MenA$_3$ neutralises toxicity by phosphorylating MenT$_3$ serine S78, which is predicted to disrupt the active site and interfere with MenT$_3$ NTase activity.

To explore this mechanism further, the web-based HADDOCK2.4 modelling software[363] was used to perform molecular docking of $MenT_3$ together with CTP and the $tRNA^{Leu}$ from *Pyrococcus horikoshii* (PDB: 1WZ2) (Figure 7.3). The resulting model, whilst only predictive, presents a possible rationale for $MenT_3$ NTase activity in transferring pyrimidines to tRNAs. The phosphodiester backbone of negatively charged tRNA molecules interacts with the electropositive $MenT_3$ groove, stretching from an NTD β-sheet and feeding through into the $MenT_3$ active site (Figure 7.3).

The importance of the conserved $MenT_3$ active site residues is also evident: $MenT_3$ D80 appears to guide the 3'-CCA motif into the active site, whilst K61, K189 and D211 cluster around CTP where they might catalyse nucleotide transfer (Figure 7.4A). The location of $MenT_3$ S78 also reaffirms its importance to antitoxicity, recessed as it is towards the rear of the active site, between the proposed binding site of both CTP and the tRNA 3'-CCA (Figure 7.4A). Indeed, an alignment of $MenT_3$ SEP to the modelled $MenT_3$:CTP:$tRNA^{Leu}$ complex shows how the phosphoserine protrudes into the active site centre where it would likely physically displace CTP or disrupt binding (Figure 7.4B). Furthermore, the presence of the phosphoserine might alter the overall surface charge, or effect minor structural changes to the binding interface which could impair coordination of the tRNA 3'-CCA with the $MenT_3$ core.

$MenT_3$ toxicity represents a novel mechanism by which a bacterial toxin inhibits the translation pathway. This is despite the large array of characterised TA systems already reported to specifically target tRNAs to inhibit protein synthesis[417,418]. That *M*. *tuberculosis* encodes at least eighty TA systems, with many of these classified as translation-inhibiting VapB-VapC systems, is intriguing[11,12]. Blocking protein synthesis would lead to a resource-sparing reduced growth state, which importantly is reversible. By possessing diverse options to target and inhibit protein synthesis, *M*. *tuberculosis* might employ TA systems as wide-ranging and condition-specific response elements to promote adaptation to hostile environments within the host[12]. Indeed, in response to various cell stress conditions, numerous TA systems have been found transcriptionally upregulated[419], suggesting, although not proving, a role in *M*. *tuberculosis* physiological responses. Alternatively, given that many *M*. *tuberculosis* TA modules are encoded on genomic islands alongside virulence and physiology genes, these systems may indirectly promote stress adaptation or pathogenicity by maintaining these regions and preventing their loss[12].

The tRNA specificity shown by $MenT_3$ is interesting but not surprising. Numerous toxins have been shown to inhibit tRNAs, usually targeting a specific tRNA or subset of tRNAs. For example, AtaT, a GNAT family toxin from *E. coli*, specifically acetylates the amine group of methionine-charged initiator $tRNA^{fMet70}$. Interestingly, a different GNAT toxin, TacT from *S*. Typhimurium, is more promiscuous, acetylating the primary amine group of multiple charged elongator tRNAs[132]. The reasons for tRNA discrimination are usually poorly understood; where characterised, specific tRNA sequence and structure motifs are usually key determinants[417]. For example, in the case of the MazF-mt9 RNase toxin from *M*. *tuberculosis*, cleavage of $tRNA^{Lys43}$ is dependent on both an intact cleavage site and appropriate secondary structure of the tRNA anticodon stem-loop[126]. Similarly, the VapC1 and VapC2 RNase toxins from *Haemophilus influenzae* require the presence of specific C-G base pairs in the $tRNA^{fMet}$ anticodon stem-loop to carry out cleavage[420].

Accordingly, investigating the selective mechanism by which $MenT_3$ discriminates among *M*. *tuberculosis* tRNAs would be a useful next step[98]. The molecular docking of $MenT_3$:CTP:$tRNA^{Leu}$ indicates an expansive region of electropositive surface charge stretching from the active site towards the NTD which appears not be involved in the modelled $tRNA^{Leu}$ interaction (Figure 7.3). This interface may simply serve to enhance the overall $MenT_3$ surface charge to facilitate general tRNA binding. Alternatively, this region might represent the presence of additional functional residues or a protein conformation better suited to facilitate the binding of different tRNAs. In addition, given the shared GNAT classification of the aforementioned AtaT and TacT toxins[70,132], yet divergent tRNA specificities, it will be interesting to elucidate the molecular mechanism of $MenT_1$ and $MenT_4$ toxicity and investigate whether any potential tRNA specificity is similarly distinct.

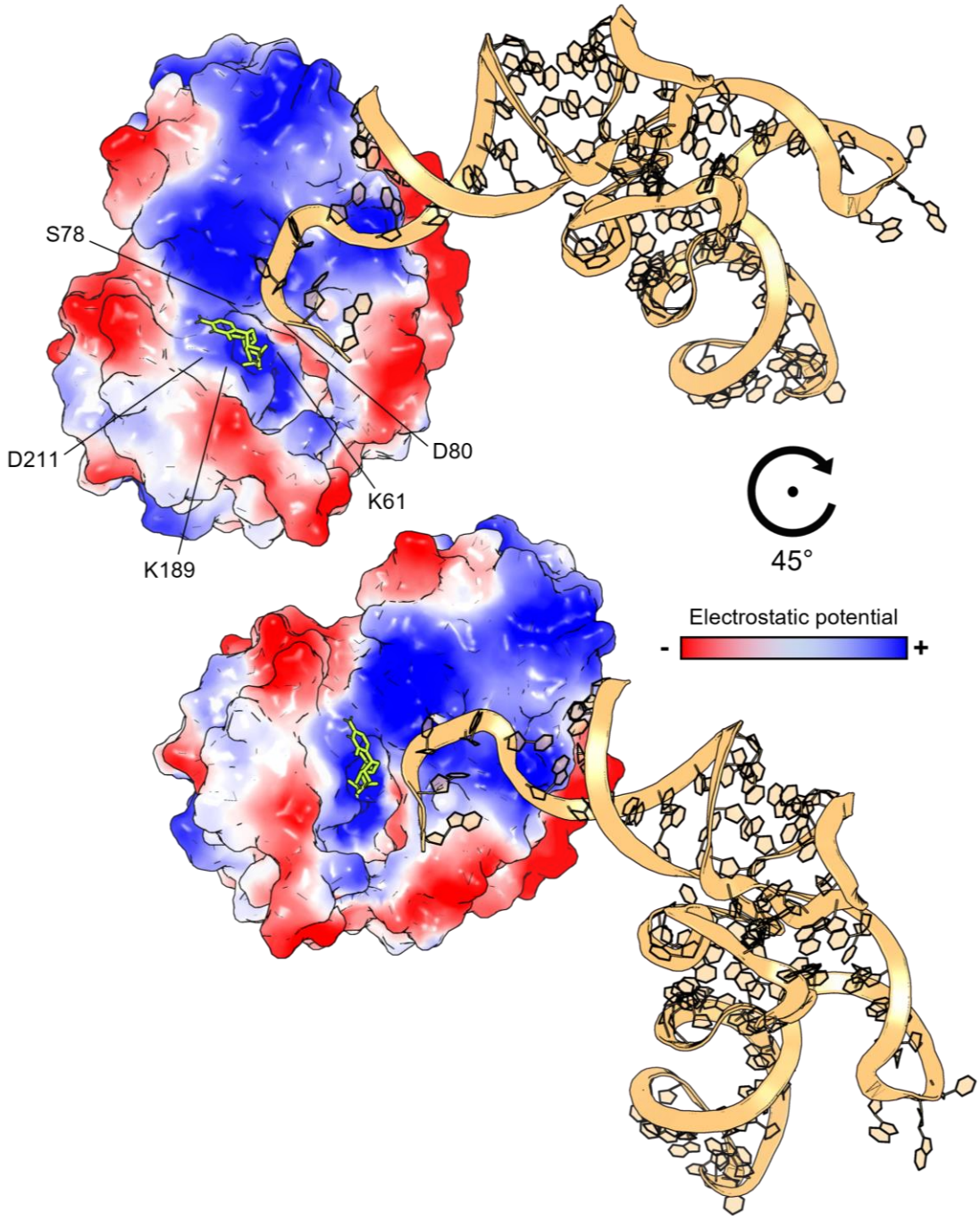


Figure 7.3. Structural modelling of the MenT$_3$:CTP:tRNA^{Leu} interface. Surface representations of MenT$_3$, coloured according to electrostatic potential, modelled in complex with CTP and tRNA^{Leu} from *P. horikoshii* (PDB: 1WZ2). Essential MenT$_3$ residues are labelled. Molecular docking was performed with HADDOCK2.4 using default settings[363]: MenT$_3$ K61, D80, K189 and D211, the tRNA^{Leu} 3'-CCA, and CTP were selected as active (interacting) residues.

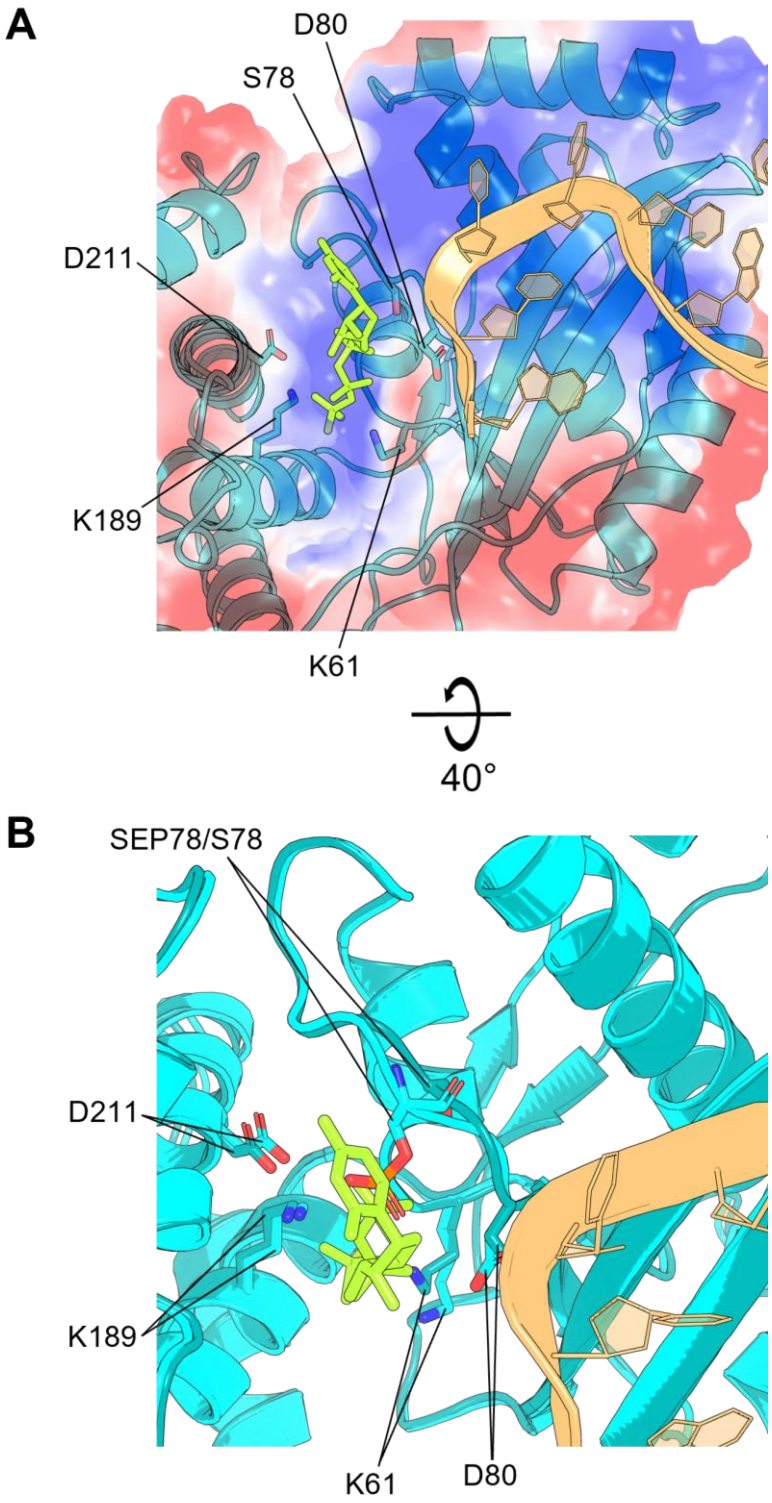


Figure 7.4. Close-up views of the MenT$_3$:CTP:tRNA^{Leu} active site interface. (A) Surface representations of MenT$_3$, coloured according to electrostatic potential, modelled in complex with CTP and tRNA^{Leu} from *P. horikoshii* (PDB: 1WZ2). The MenT$_3$ cartoon structure (teal) is visible through surface electrostatics, with essential active site residues labelled and displayed as sticks. **(B)** Alignment of MenT$_3$ SEP (aquamarine) to MenT$_3$:CTP:tRNA^{Leu}, as in (A), rotated on *x* by -40°. Essential residues are labelled and displayed as sticks. Molecular docking was performed with HADDOCK2.4 using default settings[363]: MenT$_3$ K61, D80, K189 and D211, the tRNA^{Leu} 3'-CCA, and CTP were selected as active (interacting) residues.

7.4. Interpretations and future work

Whilst the work presented here fulfilled a broad functional, structural, and biochemical characterisation of the MenA-MenT TA family, a number of additional questions were raised that require further investigation. These can be categorised into four distinct future research aims: first, definitively characterise the $MenA_2$-$MenT_2$ TA system; second, elucidate the molecular mechanism of $MenT_1$ and $MenT_4$ toxicity; third, elaborate on the antitoxic mechanisms at play in the MenA-MenT family; and fourth, investigate the functional relevance of MenA-MenT TA systems to *M*. *tuberculosis* physiology.

$MenA_2$-$MenT_2$ characterisation

Given the demonstrated activity of the $MenT_3$ and $MenT_4$ toxins in *E. coli*, the absence of any comparable TA activity from the $MenA_2$-$MenT_2$ system was surprising. Potentially, this indicates simply that $MenA_2$-$MenT_2$ is not a TA system and was misidentified. However, given its characteristic TA gene-pair architecture, as well as the predicted structural homology to both the MenA-MenT TA systems and the AbiEi-AbiEii TA module, several other explanations are possible. $MenT_2$ may be misfolded or improperly expressed in *E. coli*, resulting in the absence of any toxicity phenotypes. This could explain the lack of expression in *E. coli* from pSAT1-LIC expression constructs for both $MenT_2$ (See, *5.2*) and $MenA_2$ (Izaak Beck, personal communication). This reason could also be extended to the inactivity of $MenT_1$ in *E. coli*[98]. Alternatively, it is plausible that $MenT_1$ and $MenT_2$ are unable to detect their corresponding tRNA substrates due to modifications or the absence of the preferred target[421]. An additional possibility might be that in *E. coli*, tRNA targets are expressed at a higher level, and therefore sufficient to compensate for the inhibitory effect of $MenT_1$ and $MenT_2$ *in vivo*. Supporting this, *M*. *tuberculosis* encodes just 45 tRNA genes, whereas *E. coli* encodes 86, including four initiator tRNA genes[421,422].

Whilst the lack of activity in *E. coli* is easier to rationalise, the absence of $MenT_2$ toxicity in *M. smegmatis* is trickier to justify[98], given its closer relationship to *M*. *tuberculosis*. In this context, the $MenT_2$ target, tRNA or otherwise, may simply not be conserved outside of *M*. *tuberculosis*. *M*. *tuberculosis* is highly pathogenic and stands in stark contrast to its non-pathogenic soil-dwelling cousin *M. smegmatis*, which has separately evolved and adapted to its own distinct niche[423]. Whilst these two organisms feature conserved housekeeping genes and functions, *M. smegmatis* is unlikely to represent the same suite of virulence genes that make *M*. *tuberculosis* such an effective pathogen[424–426]. Indeed, *M*. *tuberculosis*, encoding at

least eighty TA systems[11,12], is distinct from non-pathogenic *M. smegmatis*, encoding just five[13]. The acquisition and conservation of TA systems in the MTBC has been proposed as a key contributor to the evolution and pathogenicity of *M*. *tuberculosis*[12]. This would explain the disparity of TA systems between the two organisms and offers a possible scenario where the target of the $MenT_2$ toxin is present in *M*. *tuberculosis*, but not conserved in either *M. smegmatis* or *E. coli*. As such, $MenA_2$-$MenT_2$ may yet have functional but uncharacterised TA activity within its native organism, meaning phenotypic testing in *M*. *tuberculosis* is essential in order to definitively characterise this system.

Defining $MenT_1$ and $MenT_4$ toxicity

One of the priorities for future work will be to characterise the molecular mechanism of toxicity for both $MenT_1$ and $MenT_4$. Using *in vitro* expression assays, $MenT_4$ was shown to inhibit protein synthesis, similar to $MenT_3$. Moreover, the fact that $MenT_3$ and $MenT_4$ share structural homology, where $MenT_4$ is also predicted to function as an NTase, suggests that a tRNA-targeting mechanism might be conserved. This can be investigated further by replicating the experimental characterisation of $MenT_3$ performed in the Cai *et al*. study, using nucleotide transfer assays with each of the 45 *M*. *tuberculosis* tRNAs as test substrates[98]. Interestingly, although both $MenT_3$ and $MenT_4$ inhibited protein synthesis *in vitro*, the $MenT_4$ protein concentration required for comparable inhibition was five-fold higher than $MenT_3$. This might represent that, although a fundamental mechanism is shared, $MenT_4$ has different tRNA substrate specificity to $MenT_3$.

The prospect of repeating these experiments with $MenT_1$, however, faces a specific hurdle, namely the lack of $MenT_1$ toxicity in *E. coli*. The PURExpress *in vitro* expression assay (New England Biolabs) used for characterising $MenT_3$ and $MenT_4$ toxicity consists of the purified and reconstituted components necessary for *E. coli* transcription and translation. Therefore, while as yet untested, it is unlikely that $MenT_1$ would produce similar results using this particular assay. If the mechanism of toxicity is conserved, as structural analyses suggest is possible, then nucleotide transfer assays testing *M*. *tuberculosis* tRNAs would be a valid experiment. However, to first confirm that $MenT_1$ does indeed inhibit protein synthesis, it may be necessary to independently construct a mycobacterial version of the *in vitro* transcription/translation system. In lieu of a commercial option, this would likely demand time-consuming development and optimisation. However, the proof of principle has previously been shown; Srivastava and colleagues reported the methodology for a hybrid reconstituted protein translation system drawing on a combination of *M*. *tuberculosis*

transcription factors and *M. smegmatis* tRNAs, ribosomes, and aminoacyl tRNA synthetases (AaRSs)[427]. The authors also showed that components of the mycobacterial system exhibited distinct differences to the equivalent *E. coli* PURExpress system; for example, *M. smegmatis* tRNAs were inefficiently charged by *E. coli* AaRSs, whilst the mycobacterial system was preferentially inhibited by tetracycline[427]. This might reflect structural differences in the mycobacterial translation machinery that would account for the absence of $MenT_1$ toxicity in *E. coli*, and would therefore be interesting to explore.

Clarifying MenA-MenT antitoxicity

An additional aspect of the MenA-MenT family which needs elaborating on is the question of antitoxicity. Based on current results, type II ($MenA_1$-$MenT_1$), type IV ($MenA_4$-$MenT_4$), and type VII ($MenA_3$-$MenT_3$) mechanisms are predicted. In the case of $MenA_3$-$MenT_3$, the antitoxin acts as a kinase, phosphorylating the $MenT_3$ toxin at residue S78 to neutralise toxicity[99]. However, as previously discussed, questions remain regarding whether phosphorylated $MenT_3$ retains toxicity, albeit lessened, or whether a cellular phosphatase such as PstP removes the S78 phosphoserine *in vivo* in *M. tuberculosis*[413,414] (See, *6.6*). The Ser/Thr phosphatase PstP has indeed been implicated and shown to dephosphorylate $MenT_3$ *in vitro*[99]; this relationship could be further investigated by testing the effect of PstP overexpression on $MenA_3$ antitoxicity, or carrying out *in vivo* co-purifications of PstP and $MenT_3$ to search for any detectable interactions.

The MS experiments reported in this thesis also warrant repeating to verify their results; notably, these were only performed once per sample and therefore lack the validity of multiple replicates. This would benefit the characterisation of $MenA_4$-$MenT_4$, as to date no clues to an antitoxic mechanism have been revealed. $MenA_4$ does not appear to interact with $MenT_4$ using SEC (type II), whilst no toxin modifications were detectable via MS as a result of antitoxin co-expression (type VII). Whilst this suggests a putative type IV mechanism, no definitive evidence has yet been collected that supports this hypothesis. The $MenA_4$+$MenT_4$ co-expression used for MS analyses produced notably poor protein yields, potentially indicating an unhealthy expression strain or an expression-inhibiting effect from $MenA_4$+$MenT_4$ expression (See, *5.6*). This raises questions as to the validity of the protein used. As such, the $MenA_4$+$MenT_4$ co-expression should first be repeated or optimised, prior to repetition of MS. If MS results are indeed validated, and $MenA_4$ is confirmed to not modify $MenT_4$ in a similar manner to $MenA_3$-$MenT_3$, then transcriptome profiling could instead be

carried out on cells expressing $MenA_4$, to try to identify potential antitoxin targets. Alternatively, if a type IV antitoxic mechanism is indeed responsible, then first elucidating the molecular mechanism of $MenT_4$ toxicity is key, to identify potential targets of $MenA_4$ activity.

Repeating MS experiments is especially important in the context of $MenT_1$, which was observed to potentially, and confusingly, be subject to two separate antitoxic mechanisms: co-expression of $MenA_1$ appeared to cause the putative phosphorylation of a subset of $MenT_1$, as with $MenA_3$-$MenT_3$, yet SEC experiments and the $MenA_1$:$MenT_1$ crystal structure suggested that complex formation was the cause of antitoxicity. As such, MS experiments should first be repeated using fresh protein preparations in order to verify the reproducibility of $MenA_1$-mediated $MenT_1$ modifications. If modifications are confirmed to be reproducible, kinase assays could be utilised to determine whether $MenA_1$ can phosphorylate $MenT_1$ *in vitro*.

The binding interface of $MenA_1$-$MenT_1$ should also be experimentally characterised to identify essential residues for complex formation. Structural analyses of the $MenA_1$:$MenT_1$ complex highlighted a number of residues involved in salt bridge or hydrogen bond formation between $MenA_1$ and $MenT_1$ protomers (See, *5.5.4*). $MenA_1$ residues L14 and V19 each interact with identical hydrophobic pockets in the $MenT_1$ protomers. Additionally, $MenT_1$ F38 was predicted to play an important role in complex formation, participating in a hydrophobic network around $MenA_1$ to mediate antitoxin binding, but also involved in $MenT_1$ protomer-protomer interactions. These residues should be mutated, then characterised functionally and structurally. Antitoxicity assays can test whether mutations abolish antitoxicity in *M. smegmatis*, whilst SEC protein-interaction studies can investigate whether these phenotypes are attributable to specific mutations abolishing $MenA_1$:$MenT_1$ complex formation.

Functional relevance of MenA-MenT TA systems

A final, wide-ranging aim for future work is to further expand on the mechanistic detail of $MenT_3$ toxicity and elucidate whether $MenA_3$-$MenT_3$, and by extension the other MenA-MenT TA systems, are functionally relevant to *M. tuberculosis* physiology. To begin with, functional assays could be performed to confirm whether the mechanism of growth inhibition is bacteriostatic or bactericidal. Further structural characterisation would seek to expand on modelling efforts presented in this study and resolve the crystal structure of $MenT_3$ bound to an NTP cofactor, tRNA, or potentially both. This would further elaborate on

the molecular mechanism by which $MenT_3$ binds these ligands and provide mechanistic insight into the conditions and essential residues required for toxicity.

Following this, any potential physiological roles could next be investigated. A previous TraSH study showed that the *menA₃* gene could not be disrupted and is essential for *M*. *tuberculosis* growth[335], suggesting that $MenT_3$ is highly toxic *in vivo*. Separately, we reported in the Cai *et al*. study that deletion of the entire *menA₃-menT₃* locus from *M*. *tuberculosis* H37Rv could not be achieved in one step, which was putatively linked to toxicity from residual $MenT_3$[98]. Taken together, these reports suggest that the chromosomal expression of $MenT_3$ is more than sufficient to inhibit growth in *M*. *tuberculosis*. This potent toxicity, even from innate expression levels, highlights the requirement for effective and constant regulation by $MenA_3$, and suggests a possible physiological role for this system in *M*. *tuberculosis*.

To characterise any potential role, a wide range of experiments could be performed *in vivo* in *M*. *tuberculosis*, though due to the time and biosecurity restrictions inherent to *M*. *tuberculosis* research, these would be labour intensive and long-term goals. RNA-seq experiments could be used to assess *menA₃-menT₃* transcript levels following exposure to a variety of cell stress conditions; this could offer clues as to whether $MenA_3$-$MenT_3$ is transcriptionally activated in response to specific stresses. Similarly, the growth of WT *M*. *tuberculosis* H37Rv and a Δ*menA₃-menT₃* strain could be monitored and compared during and after exposure to cell stresses, to assess for immediate and delayed phenotypic changes that might be caused by the loss of $MenA_3$-$MenT_3$. *M. tuberculosis* is typically resident in macrophages during chronic host infections, therefore comparing the survivability of WT and Δ*menA₃-menT₃* strains following murine macrophage internalisation would be valuable. These experiments could also be expanded to include chromosomal mutants of the characterised $MenT_3$ active site residues, to confirm their essentiality to $MenT_3$ toxicity *in vivo* and assess whether mutations cause the abolition of any potential $MenA_3$-$MenT_3$-related growth phenotypes.

Furthermore, given that $MenT_3$ was shown to target and prevent tRNA aminoacylation *in vitro*[98], it would also be interesting to assess the downstream effects on global translation *in vivo*. Although speculative, a potential consequence of $MenT_3$ activity might be the accumulation of deacylated tRNAs at the ribosomal A site. This could in turn cause the RelA-dependent synthesis of (p)ppGpp, leading to an SR-like regulatory cascade which, among other consequences, inhibits translation (See, *1.3.3*). Therefore, assessing changes to global translation and protein synthesis rates due to $MenT_3$ expression could implicate its activity

in biologically relevant processes. This could initially be performed using a plasmid-based expression system, in *E. coli* WT and Δ*relA* strains, before later experiments could also compare changes in global translation between *M*. *tuberculosis* WT and $\Delta menA_3$-$menT_3$ strains.

Considering the reportedly strong toxicity resulting from chromosomal $MenT_3$ expression and the essentiality of $MenA_3$ to *M*. *tuberculosis* growth[98,335], it would be interesting to assess whether cellular proteases in *M*. *tuberculosis* recognise $MenA_3$ for degradation. Significantly, the results presented in this thesis indicate that $MenA_3$ does not form a T:A complex with $MenT_3$, implying that $MenA_3$ would not benefit from the same protection against proteolysis that T:A complexes are proposed to confer[20,198]. A number of *M*. *tuberculosis* antitoxins have been reported to be putative protease substrates, including MazE10, VapB20, VapB47, and RelB1 (ClpC1P1P2), and HigA1 (ClpX)[428]. These relationships have commonly, if contentiously, been attributed to the bacterial stress response, where stress-activated proteases are proposed to cause antitoxin degradation and toxin liberation, leading to growth inhibition and cell survival (See, *1.6*). Testing for such a relationship would require the construction of various *M*. *tuberculosis* protease deletion strains, which could then be screened for changes in $MenA_3$ transcript levels following treatment with, for example, chloramphenicol to inhibit translation[21]. This could provide greater insight into the regulatory mechanisms underpinning the $MenA_3$-$MenT_3$ system and provide clues to possible biological roles or stress responses involving this system.

Lastly, the AbiEi-AbiEii TA module from *S. agalactiae* functions as an Abi system to protect the cell from phage infection. Due to structural homology shared by AbiEi-AbiEii and MenA-MenT systems, as well as similar mechanisms of regulation and toxin activity, initial efforts were begun to test whether the MenA-MenT family shares this protective role. However, at the time of writing, these experiments had not progressed beyond the cloning stage. It would therefore be interesting to characterise the MenA-MenT systems in this context and determine if an Abi function is also conserved.

7.5. Conclusion

At its outset, this project aimed to comprehensively characterise the novel MenA-MenT family of TA systems from *M*. *tuberculosis*. The work presented in this thesis describes how these aims were broadly achieved. Two of these systems were shown to be functioning TA modules in *E. coli*. The crystal structures of three MenT toxins were solved for the first time, which highlighted structural conservation and led to the characterisation of a conserved MenT active site. Protein interaction studies and autoregulation assays revealed that despite its limited size, the MenA-MenT family consists of a variety of TA types, including putative type II and type IV systems, and the newly classified type VII system. Finally, two of these systems were shown to inhibit protein synthesis *in vitro*, which was separately linked to the NTase activity predicted through structural analyses.

The results described in this thesis enhance our working knowledge of the broad repertoire of TA systems encoded by *M*. *tuberculosis*, revealing a novel TA family which may be considerably more widespread than is currently understood. The diverse antitoxic mechanisms that feature within the MenA-MenT family juxtaposes with the predicted structural and mechanistic conservation of the toxins, making them a fascinating prospect for study. Elucidating their relative contribution to *M*. *tuberculosis* physiology presents an exciting challenge, and the experimental characterisation and prospective research outlined in this study will help guide the direction of future work.

References

1. Ogura, T. & Hiraga, S. Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4784–4788 (1983).
2. Gerdes, K., Rasmussen, P. B. & Molin, S. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3116–3120 (1986).
3. Gotfredsen, M. & Gerdes, K. The *Escherichia coli relBE* genes belong to a new toxin–antitoxin gene family. *Mol. Microbiol.* **29**, 1065–1076 (1998).
4. Masuda, Y., Miyakawa, K., Nishimura, Y. & Ohtsubo, E. *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. *J. Bacteriol.* **175**, 6850 (1993).
5. Lehnherr, H., Maguin, E., Jafri, S. & Yarmolinsky, M. B. Plasmid Addiction Genes of Bacteriophage P1: *doc*, which Causes Cell Death on Curing of Prophage, and *phd*, which Prevents Host Death when Prophage is Retained. *J. Mol. Biol.* **233**, 414–428 (1993).
6. Szekeres, S., Dauti, M., Wilde, C., Mazel, D. & Rowe-Magnus, D. A. Chromosomal toxin–antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol. Microbiol.* **63**, 1588–1605 (2007).
7. Page, R. & Peti, W. Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.* **12**, 208–214 (2016).
8. Pandey, D. P. & Gerdes, K. Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.* **33**, 966–976 (2005).
9. Yamaguchi, Y., Park, J.-H. & Inouye, M. Toxin-Antitoxin Systems in Bacteria and Archaea. *Annu. Rev. Genet.* **45**, 61–79 (2011).
10. Kim, J.-S. & Wood, T. K. Persistent Persister Misperceptions. *Front. Microbiol.* **7**, 2134 (2016).
11. Sala, A., Bordes, P. & Genevaux, P. Multiple Toxin-Antitoxin Systems in *Mycobacterium tuberculosis*. *Toxins (Basel).* **6**, 1002–1020 (2014).
12. Ramage, H. R., Connolly, L. E. & Cox, J. S. Comprehensive Functional Analysis of *Mycobacterium tuberculosis* Toxin-Antitoxin Systems: Implications for Pathogenesis, Stress Responses, and Evolution. *PLOS Genet.* **5**, e1000767 (2009).
13. Shao, Y. *et al.* TADB: a web-based resource for Type 2 toxin–antitoxin loci in bacteria and archaea. *Nucleic Acids Res.* **39**, D606–D611 (2011).
14. Deter, H. S., Jensen, R. V., Mather, W. H. & Butzin, N. C. Mechanisms for Differential Protein Production in Toxin–Antitoxin Systems. *Toxins (Basel).* **9**, 211 (2017).
15. Gerdes, K., Christensen, S. K. & Løbner-Olesen, A. Prokaryotic toxin–antitoxin stress

response loci. *Nat. Rev. Microbiol.* **3**, 371–382 (2005).

16. Magnuson, R. D. Hypothetical functions of toxin-antitoxin systems. *J. Bacteriol.* **189**, 6089–92 (2007).
17. Hayes, F. & Van Melderen, L. Toxins-antitoxins: diversity, evolution and function. *Crit. Rev. Biochem. Mol. Biol.* **46**, 386–408 (2011).
18. Muthuramalingam, M., White, J. C. & Bourne, C. R. Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways. *Toxins (Basel).* **8**, 1–16 (2016).
19. Schuster, C. F. & Bertram, R. Toxin–antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS Microbiol. Lett.* **340**, 73–85 (2013).
20. Song, S. & Wood, T. K. Toxin/Antitoxin System Paradigms: Toxins Bound to Antitoxins Are Not Likely Activated by Preferential Antitoxin Degradation. *Adv. Biosyst.* **4**, 1900290 (2020).
21. LeRoux, M., Culviner, P. H., Liu, Y. J., Littlehale, M. L. & Laub, M. T. Stress Can Induce Transcription of Toxin-Antitoxin Systems without Activating Toxin. *Mol. Cell* **79**, 280-292.e8 (2020).
22. Jaffe, Aline; Ogura, Teru; Hiraga, S. Effects of the *ccd* Function of the F Plasmid on Bacterial Growth. *J. Bacteriol.* **163**, 841–849 (1985).
23. Hiraga, S., Jaffé, A., Ogura, T., Mori, H. & Takahashi, H. F plasmid *ccd* mechanism in *Escherichia coli*. *J. Bacteriol.* **166**, 100 (1986).
24. Van Melderen, L., Bernard, P. & Couturier, M. Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol. Microbiol.* **11**, 1151–1157 (1994).
25. Van Melderen, L. *et al.* ATP-dependent Degradation of CcdA by Lon Protease. Effects of Secondary Structure and Heterologous Interactions*. *J. Biol. Chem.* **271**, 27730–27738 (1996).
26. Maki, S., Takiguchi, S., Miki, T. & Horiuchi, T. Modulation of DNA supercoiling activity of *Escherichia coli* DNA gyrase by F plasmid proteins: Antagonistic actions of LetA (CcdA) and LetD (CcdB) proteins*. *J. Biol. Chem.* **267**, 12244–12251 (1992).
27. Bernard, P. *et al.* The F Plasmid CcdB Protein Induces Efficient ATP-dependent DNA Cleavage by Gyrase. *J. Mol. Biol.* **234**, 534–541 (1993).
28. Miki, T., Ae Park, J., Nagao, K., Murayama, N. & Horiuchi, T. Control of segregation of chromosomal DNA by sex factor F in *Escherichia coli*: Mutants of DNA gyrase subunit A suppress *letD* (*ccdB*) product growth inhibition. *J. Mol. Biol.* **225**, 39–52 (1992).
29. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves

poisoning of DNA-topoisomerase II complexes. *J. Mol. Biol.* **226**, 735–745 (1992).

30. Gerdes, K. *et al.* Mechanism of postsegregational killing by the *hok* gene product of the *parB* system of plasmid R1 and its homology with the *relF* gene product of the *E. coli relB* operon. *EMBO J.* **5**, 2023–2029 (1986).
31. Fineran, P. C. *et al.* The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 894–899 (2009).
32. Pecota, D. C. & Wood, T. K. Exclusion of T4 phage by the *hok*/*sok* killer locus from plasmid R1. *J. Bacteriol.* **178**, 2044–2050 (1996).
33. Lewis, K. Persister Cells. *Annu. Rev. Microbiol.* **64**, 357–372 (2010).
34. Ronneau, S. & Helaine, S. Clarifying the Link between Toxin–Antitoxin Modules and Bacterial Persistence. *J. Mol. Biol.* **431**, 3462–3471 (2019).
35. Gerdes, K., Gultyaev, A. P., Franch, T., Pedersen, K. & Mikkelsen, N. D. Antisense RNA-Regulated Programmed Cell Death. *Annu. Rev. Genet.* **31**, 1–31 (1997).
36. Gerdes, K., Helin, K., Christensen, O. W. & Løbner-Olesen, A. Translational control and differential RNA decay are key elements regulating postsegregational expression of the killer protein encoded by the *parB* locus of plasmid R1. *J. Mol. Biol.* **203**, 119–129 (1988).
37. Thisted, T. & Gerdes, K. Mechanism of post-segregational killing by the *hok*/*sok* system of plasmid R1: Sok antisense RNA regulates *hok* gene expression indirectly through the overlapping *mok* gene. *J. Mol. Biol.* **223**, 41–54 (1992).
38. Gerdes, K., Thisted, T. & Martinussen, J. Mechanism of post-segregational killing by the *hok*/*sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Mol. Microbiol.* **4**, 1807–1818 (1990).
39. Franch, T., Gultyaev, A. P. & Gerdes, K. Programmed cell death by *hok*/*sok* of plasmid R1: Processing at the *hok* mRNA 3′-end triggers structural rearrangements that allow translation and antisense RNA binding. *J. Mol. Biol.* **273**, 38–51 (1997).
40. Thisted, T., Sørensen, N. S. & Gerdes, K. Mechanism of Post-segregational Killing: Secondary Structure Analysis of the Entire Hok mRNA from Plasmid R1 Suggests a Fold-back Structure that Prevents Translation and Antisense RNA Binding. *J. Mol. Biol.* **247**, 859–873 (1995).
41. Thisted, T., Nielsen, A. K. & Gerdes, K. Mechanism of post-segregational killing: translation of Hok, SrnB and Pnd mRNAs of plasmids R1, F and R483 is activated by 3′-end processing. *EMBO J.* **13**, 1950–1959 (1994).
42. Gerdes, K., Nielsen, A., Thorsted, P. & Wagner, E. G. H. Mechanism of Killer Gene

Activation. Antisense RNA-dependent RNase III Cleavage Ensures Rapid Turn-over of the Stable Hok, SrnB and PndA Effector Messenger RNAs. *J. Mol. Biol.* **226**, 637–649 (1992).

43. Thisted, T., Sørensen, N. S., Wagner, E. G. & Gerdes, K. Mechanism of post-segregational killing: Sok antisense RNA interacts with Hok mRNA via its 5'-end single-stranded leader and competes with the 3'-end of Hok mRNA for binding to the *mok* translational initiation region. *EMBO J.* **13**, 1960–1968 (1994).
44. Durand, S., Jahn, N., Condon, C. & Brantl, S. Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.* **9**, 1491–1497 (2012).
45. Fozo, E. M. *et al.* Abundance of type I toxin–antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res.* **38**, 3743–3759 (2010).
46. Peltier, J. *et al.* Type I toxin-antitoxin systems contribute to the maintenance of mobile genetic elements in *Clostridioides difficile*. *Commun. Biol.* **3**, 1–13 (2020).
47. Pedersen, K. & Gerdes, K. Multiple *hok* genes on the chromosome of *Escherichia coli*. *Mol. Microbiol.* **32**, 1090–1102 (1999).
48. Faridani, O. R., Nikravesh, A., Pandey, D. P., Gerdes, K. & Good, L. Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. *Nucleic Acids Res.* **34**, 5915 (2006).
49. Fozo, E. M., Hemm, M. R. & Storz, G. Small Toxic Proteins and the Antisense RNAs That Repress Them. *Microbiol. Mol. Biol. Rev.* **72**, 579–589 (2008).
50. Greenfield, T. J. & Weaver, K. E. Antisense RNA regulation of the pAD1 *par* post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Mol. Microbiol.* **37**, 661–670 (2000).
51. Weaver, K. E. & Tritle, D. J. Identification and Characterization of an *Enterococcus faecalis* Plasmid pAD1-Encoded Stability Determinant Which Produces Two Small RNA Molecules Necessary for Its Function. *Plasmid* **32**, 168–181 (1994).
52. Weaver, K. E. *et al.* Identification and characterization of a family of toxin–antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 *par* addiction module. *Microbiology* **155**, 2930 (2009).
53. Silvaggi, J. M., Perkins, J. B. & Losick, R. Small Untranslated RNA Antitoxin in *Bacillus subtilis*. *J. Bacteriol.* **187**, 6641 (2005).
54. Vogel, J., Argaman, L., Wagner, E. G. H. & Altuvia, S. The Small RNA IstR Inhibits Synthesis of an SOS-Induced Toxic Peptide. *Curr. Biol.* **14**, 2271–2276 (2004).
55. Fozo, E. M. *et al.* Repression of small toxic protein synthesis by the Sib and OhsC small

RNAs. *Mol. Microbiol.* **70**, 1076–1093 (2008).

56. Kawano, M., Oshima, T., Kasai, H. & Mori, H. Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a *cis*-encoded small antisense RNA in *Escherichia coli*. *Mol. Microbiol.* **45**, 333–349 (2002).
57. Kawano, M., Aravind, L. & Storz, G. An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol. Microbiol.* **64**, 738 (2007).
58. Guo, Y. *et al.* RalR (a DNase) and RalA (a small RNA) form a type I toxin–antitoxin system in *Escherichia coli*. *Nucleic Acids Res.* **42**, 6448–6462 (2014).
59. Coray, D. S., Wheeler, N. E., Heinemann, J. A. & Gardner, P. P. Why so narrow: Distribution of anti-sense regulated, type I toxin-antitoxin systems compared with type II and type III systems. *RNA Biol.* **14**, 275–280 (2017).
60. Kawano, M., Reynolds, A. A., Miranda-Rios, J. & Storz, G. Detection of 5′- and 3′-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* **33**, 1040 (2005).
61. Zhang, S.-P. *et al.* Type II toxin–antitoxin system in bacteria: activation, function, and mode of action. *Biophys. Reports* **6**, 68–79 (2020).
62. Khoo, S. K. *et al.* Molecular and Structural Characterization of the PezAT Chromosomal Toxin-Antitoxin System of the Human Pathogen *Streptococcus pneumoniae**. *J. Biol. Chem.* **282**, 19606–19618 (2007).
63. Schumacher, M. A. *et al.* Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science.* **323**, 396–401 (2009).
64. Dao-Thi, M. H. *et al.* Molecular Basis of Gyrase Poisoning by the Addiction Toxin CcdB. *J. Mol. Biol.* **348**, 1091–1102 (2005).
65. Harms, A., Brodersen, D. E., Mitarai, N. & Gerdes, K. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol. Cell* **70**, 768–784 (2018).
66. Masuda, H. & Inouye, M. Toxins of Prokaryotic Toxin-Antitoxin Systems with Sequence-Specific Endoribonuclease Activity. *Toxins (Basel).* **9**, 140 (2017).
67. Jiang, Y., Pogliano, J., Helinski, D. R. & Konieczny, I. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Mol. Microbiol.* **44**, 971–979 (2002).
68. Castro-Roa, D. *et al.* The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat. Chem. Biol.* **9**, 811–817 (2013).
69. Harms, A. *et al.* Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Rep.* **12**, 1497–1507 (2015).

70. Jurėnas, D. *et al.* AtaT blocks translation initiation by N-acetylation of the initiator $tRNA^{fMet}$. *Nat. Chem. Biol.* **13**, 640–646 (2017).

71. Zhang, Y. *et al.* MazF Cleaves Cellular mRNAs Specifically at ACA to Block Protein Synthesis in *Escherichia coli*. *Mol. Cell* **12**, 913–923 (2003).

72. Goeders, N., Drèze, P.-L. & Van Melderen, L. Relaxed Cleavage Specificity within the RelE Toxin Family. *J. Bacteriol.* **195**, 2541–2549 (2013).

73. Blower, T. R. *et al.* Mutagenesis and Functional Characterization of the RNA and Protein Components of the *toxIN* Abortive Infection and Toxin-Antitoxin Locus of *Erwinia*. *J. Bacteriol.* **191**, 6029–6039 (2009).

74. Blower, T. R. *et al.* A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat. Struct. Mol. Biol.* **18**, 185–190 (2011).

75. Goeders, N., Chai, R., Chen, B., Day, A. & Salmond, G. P. C. Structure, Evolution, and Functions of Bacterial Type III Toxin-Antitoxin Systems. *Toxins (Basel).* **8**, (2016).

76. Short, F. L. *et al.* Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E241–E249 (2013).

77. Short, F. L., Akusobi, C., Broadhurst, W. R. & Salmond, G. P. C. The bacterial Type III toxin-antitoxin system, ToxIN, is a dynamic protein-RNA complex with stability-dependent antiviral abortive infection activity. *Sci. Rep.* **8**, (2018).

78. Blower, T. R. *et al.* Identification and classification of bacterial Type III toxin–antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res.* **40**, 6158 (2012).

79. Short, F. L., Monson, R. E. & Salmond, G. P. C. A Type III protein-RNA toxin-antitoxin system from *Bacillus thuringiensis* promotes plasmid retention during spore development. *RNA Biol.* **12**, 933 (2015).

80. Rao, F. *et al.* Co-evolution of quaternary organization and novel RNA tertiary interactions revealed in the crystal structure of a bacterial protein–RNA toxin–antitoxin system. *Nucleic Acids Res.* **43**, 9529 (2015).

81. Masuda, H., Tan, Q., Awano, N., Wu, K.-P. & Inouye, M. YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. *Mol. Microbiol.* **84**, 979–989 (2012).

82. Tan, Q., Awano, N. & Inouye, M. YeeV is an *Escherichia coli* toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. *Mol. Microbiol.* **79**, 109–118 (2011).

83. Dy, R. L., Przybilski, R., Semeijn, K., Salmond, G. P. C. & Fineran, P. C. A widespread

bacteriophage abortive infection system functions through a Type IV toxin–antitoxin mechanism. *Nucleic Acids Res.* **42**, 4590–4605 (2014).

84. Hampton, H. G. *et al.* AbiEi Binds Cooperatively to the Type IV *abiE* Toxin–Antitoxin Operator Via a Positively-Charged Surface and Causes DNA Bending and Negative Autoregulation. *J. Mol. Biol.* **430**, 1141–1156 (2018).
85. Aakre, C. D., Phung, T. N., Huang, D. & Laub, M. T. A Bacterial Toxin Inhibits DNA Replication Elongation through a Direct Interaction with the β Sliding Clamp. *Mol. Cell* **52**, 617–628 (2013).
86. Wang, X. *et al.* A Novel Type V TA System Where mRNA for Toxin GhoT is Cleaved by Antitoxin GhoS. *Nat. Chem. Biol.* **8**, 855 (2012).
87. Beloglazova, N. *et al.* A Novel Family of Sequence-specific Endoribonucleases Associated with the Clustered Regularly Interspaced Short Palindromic Repeats. *J. Biol. Chem.* **283**, 20361 (2008).
88. Wang, X. *et al.* Type II toxin/antitoxin MqsR/MqsA controls type V toxin/antitoxin GhoT/GhoS. *Environ. Microbiol.* **15**, 1734–1744 (2013).
89. Yamaguchi, Y., Park, J.-H. & Inouye, M. MqsR, a Crucial Regulator for Quorum Sensing and Biofilm Formation, Is a GCU-specific mRNA Interferase in *Escherichia coli*. *J. Biol. Chem.* **284**, 28746–28753 (2009).
90. Johnson, A. & O'Donnell, M. Cellular DNA Replicases: Components and Dynamics at the Replication Fork. *Annu. Rev. Biochem.* **74**, 283–315 (2005).
91. Saro, F. J. L. de, Georgescu, R. E., Goodman, M. F. & O'Donnell, M. Competitive processivity-clamp usage by DNA polymerases during DNA replication and repair. *EMBO J.* **22**, 6408 (2003).
92. Saro, F. J. L. de, Marinus, M. G., Modrich, P. & O'Donnell, M. The β Sliding Clamp Binds to Multiple Sites within MutL and MutS*. *J. Biol. Chem.* **281**, 14340–14349 (2006).
93. Kurz, M., Dalrymple, B., Wijffels, G. & Kongsuwan, K. Interaction of the sliding clamp β-subunit and Hda, a DnaA-related protein. *J. Bacteriol.* **186**, 3508–3515 (2004).
94. Wang, X., Yao, J., Sun, Y.-C. & Wood, T. K. Type VII Toxin/Antitoxin Classification System for Antitoxins that Enzymatically Neutralize Toxins. *Trends Microbiol.* **29**, 388–393 (2021).
95. García-Contreras, R., Zhang, X.-S., Kim, Y. & Wood, T. K. Protein Translation and Cell Death: The Role of Rare tRNAs in Biofilm Formation and in Activating Dormant Phage Killer Genes. *PLoS One* **3**, e2394 (2008).
96. Marimon, O. *et al.* An oxygen-sensitive toxin–antitoxin system. *Nat. Commun.* **7**, 1–10 (2016).

97. Yao, J. *et al.* Novel polyadenylylation-dependent neutralization mechanism of the HEPN/MNT toxin/antitoxin system. *Nucleic Acids Res.* **48**, 11054–11067 (2020).

98. Cai, Y. *et al.* A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems. *Sci. Adv.* **6**, eabb6651 (2020).

99. Yu, X. *et al.* Characterization of a toxin-antitoxin system in *Mycobacterium tuberculosis* suggests neutralization by phosphorylation as the antitoxicity mechanism. *Commun. Biol.* **3**, 1–15 (2020).

100. Barrios, A. F. G., Zuo, R., Ren, D. & Wood, T. K. Hha, YbaJ, and OmpA regulate *Escherichia coli* K12 biofilm formation and conjugation plasmids abolish motility. *Biotechnol. Bioeng.* **93**, 188–200 (2006).

101. Prüß, B. M., Besemann, C., Denton, A. & Wolfe, A. J. A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. *J. Bacteriol.* **188**, 3731–3739 (2006).

102. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V. & Aravind, L. Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct* **8**, 1–28 (2013).

103. Jia, X. *et al.* Structure–function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN–MNT toxin–antitoxin system. *J. Biol. Chem.* **293**, 6812–6823 (2018).

104. Yao, J. *et al.* Identification and characterization of a HEPN-MNT family type II toxin–antitoxin in *Shewanella oneidensis*. *Microb. Biotechnol.* **8**, 961–973 (2015).

105. Jurėnas, D., Garcia-Pino, A. & Van Melderen, L. Novel toxins from type II toxin-antitoxin systems with acetyltransferase activity. *Plasmid* **93**, 30–35 (2017).

106. Guglielmini, J. & Van Melderen, L. Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mob. Genet. Elements* **1**, 283–306 (2011).

107. Easter, C. L., Sobecky, P. A. & Helinski, D. R. Contribution of different segments of the *par* region to stable maintenance of the broad-host-range plasmid RK2. *J. Bacteriol.* **179**, 6472–6479 (1997).

108. Roberts, R. C., Ström, A. R. & Helinski, D. R. The *parDE* Operon of the Broad-host-range Plasmid RK2 Specifies Growth Inhibition Associated with Plasmid Loss. *J. Mol. Biol.* **237**, 35–51 (1994).

109. Maki, S., Takiguchi, S., Horiuchi, T., Sekimizu, K. & Miki, T. Partner Switching Mechanisms in Inactivation and Rejuvenation of *Escherichia coli* DNA Gyrase by F

Plasmid Proteins LetD (CcdB) and LetA (CcdA). *J. Mol. Biol.* **256**, 473–482 (1996).

110. Critchlow, S. E. *et al.* The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: induction of DNA cleavage and blocking of transcription. *J. Mol. Biol.* **273**, 826–839 (1997).
111. Yuan, J. *et al. Vibrio cholerae* ParE2 Poisons DNA Gyrase via a Mechanism Distinct from Other Gyrase Inhibitors. *J. Biol. Chem.* **285**, 40397–40408 (2010).
112. Jankevicius, G., Ariza, A., Ahel, M. & Ahel, I. The Toxin-Antitoxin System DarTG Catalyzes Reversible ADP-Ribosylation of DNA. *Mol. Cell* **64**, 1109 (2016).
113. Schuller, M. *et al.* Molecular basis for DarT ADP-ribosylation of a DNA base. *Nature* **596**, 597–602 (2021).
114. Heller, D. M., Tavag, M. & Hochschild, A. CbtA toxin of *Escherichia coli* inhibits cell division and cell elongation via direct and independent interactions with FtsZ and MreB. *PLoS Genet.* **13**, (2017).
115. Goeders, N. & Van Melderen, L. Toxin-Antitoxin Systems as Multilevel Interaction Systems. *Toxins (Basel).* **6**, 304–324 (2014).
116. Poulsen, L. K., Refn, A., Molin, S. & Andersson, P. Topographic analysis of the toxic Gef protein from *Escherichia coli*. *Mol. Microbiol.* **5**, 1627–1637 (1991).
117. Unoson, C. & Wagner, E. G. H. A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol. Microbiol.* **70**, 258–270 (2008).
118. Weel-Sneve, R. *et al.* Single Transmembrane Peptide DinQ Modulates Membrane-Dependent Activities. *PLOS Genet.* **9**, e1003260 (2013).
119. Ono, T., Akimoto, S., Ono, K. & Ohnishi, Y. Plasmid genes increase membrane permeability in *Escherichia coli*. *Biochim. Biophys. Acta* **867**, 81–88 (1986).
120. Weaver, K. E. *et al. Enterococcus faecalis* plasmid pAD1-encoded Fst toxin affects membrane permeability and alters cellular responses to lantibiotics. *J. Bacteriol.* **185**, 2169–2177 (2003).
121. Mutschler, H., Gebhardt, M., Shoeman, R. L. & Meinhart, A. A Novel Mechanism of Programmed Cell Death in Bacteria by Toxin–Antitoxin Systems Corrupts Peptidoglycan Synthesis. *PLOS Biol.* **9**, e1001033 (2011).
122. Yamaguchi, Y. & Inouye, M. Chapter 12: mRNA Interferases, Sequence-Specific Endoribonucleases from the Toxin–Antitoxin Systems. *Prog. Mol. Biol. Transl. Sci.* **85**, 467–500 (2009).
123. Pedersen, K. *et al.* The Bacterial Toxin RelE Displays Codon-Specific Cleavage of mRNAs in the Ribosomal A Site. *Cell* **112**, 131–140 (2003).
124. Zhang, Y. & Inouye, M. The Inhibitory Mechanism of Protein Synthesis by YoeB, an

Escherichia coli Toxin*. *J. Biol. Chem.* **284**, 6627–6638 (2009).

125. Zhang, Y. & Inouye, M. RatA (YfjG), an *Escherichia coli* toxin, inhibits 70S ribosome association to block translation initiation. *Mol. Microbiol.* **79**, 1418–1429 (2011).
126. Schifano, J. M. *et al.* tRNA is a new target for cleavage by a MazF toxin. *Nucleic Acids Res.* **44**, 1256–1270 (2016).
127. Winther, K. S., Tree, J. J., Tollervey, D. & Gerdes, K. VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. *Nucleic Acids Res.* **44**, 9860–9871 (2016).
128. Cruz, J. W. *et al.* Growth-regulating *Mycobacterium tuberculosis* VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat. Commun.* **6**, 1–12 (2015).
129. Winther, K. S., Brodersen, D. E., Brown, A. K. & Gerdes, K. VapC20 of *Mycobacterium tuberculosis* cleaves the Sarcin–Ricin loop of 23S rRNA. *Nat. Commun.* **4**, 1–9 (2013).
130. Winther, K. S. & Gerdes, K. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7403–7407 (2011).
131. Lopes, A. P. Y. *et al.* VapC from the Leptospiral VapBC Toxin-Antitoxin Module Displays Ribonuclease Activity on the Initiator tRNA. *PLoS One* **9**, e101678 (2014).
132. Cheverton, A. M. *et al.* A *Salmonella* Toxin Promotes Persister Formation through Acetylation of tRNA. *Mol. Cell* **63**, 86–96 (2016).
133. Rycroft, J. A. *et al.* Activity of acetyltransferase toxins involved in *Salmonella* persister formation during macrophage infection. *Nat. Commun.* **9**, 1–11 (2018).
134. Comeau, A. M. *et al.* Exploring the prokaryotic virosphere. *Res. Microbiol.* **159**, 306–313 (2008).
135. Shabbir, M. A. B. *et al.* Bacteria vs. Bacteriophages: Parallel Evolution of Immune Arsenals. *Front. Microbiol.* **7**, (2016).
136. Lopatina, A., Tal, N. & Sorek, R. Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).
137. Yu, Y. T. & Snyder, L. Translation elongation factor Tu cleaved by a phage-exclusion system. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 802–806 (1994).
138. Parma, D. H. *et al.* The Rex system of bacteriophage lambda: tolerance and altruistic cell death. *Genes Dev.* **6**, 497–510 (1992).
139. Emond, E. *et al.* AbiQ, an abortive infection mechanism from *Lactococcus lactis*. *Appl. Environ. Microbiol.* **64**, 4748–4756 (1998).
140. Samson, J. E., Spinelli, S., Cambillau, C. & Moineau, S. Structure and activity of AbiQ, a lactococcal endoribonuclease belonging to the type III toxin–antitoxin system. *Mol.*

Microbiol. **87**, 756–768 (2013).

141. Chopin, M. C., Chopin, A. & Bidnenko, E. Phage abortive infection in lactococci: variations on a theme. *Curr. Opin. Microbiol.* **8**, 473–479 (2005).
142. Tangney, M. & Fitzgerald, G. F. Effectiveness of the lactococcal abortive infection systems AbiA, AbiE, AbiF and AbiG against P335 type phages. *FEMS Microbiol. Lett.* **210**, 67–72 (2002).
143. Garvey, P., Fitzgerald, G. F. & Hill, C. Cloning and DNA Sequence Analysis of Two Abortive Infection Phage Resistance Determinants from the Lactococcal Plasmid pNP40. *Appl. Environ. Microbiol.* **61**, 4321–4328 (1995).
144. Peeters, S. H. & de Jonge, M. I. For the greater good: Programmed cell death in bacterial communities. *Microbiol. Res.* **207**, 161–169 (2018).
145. Slavcev, R. A. & Hayes, S. Stationary phase-like properties of the bacteriophage λ Rex exclusion phenotype. *Mol. Genet. Genomics* **269**, 40–48 (2003).
146. Stewart, P. S. & Franklin, M. J. Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* **6**, 199–210 (2008).
147. Flemming, H.-C. *et al.* Biofilms: an emergent form of bacterial life. *Nat. Rev. Microbiol.* **14**, 563–575 (2016).
148. Hall-Stoodley, L., Costerton, J. W. & Stoodley, P. Bacterial biofilms: from the Natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2**, 95–108 (2004).
149. Vestby, L. K., Grønseth, T., Simm, R. & Nesse, L. L. Bacterial Biofilm and its Role in the Pathogenesis of Disease. *Antibiotics* **9**, (2020).
150. Ren, D., Bedzyk, L. A., Thomas, S. M., Ye, R. W. & Wood, T. K. Gene expression in *Escherichia coli* biofilms. *Appl. Microbiol. Biotechnol.* **64**, 515–524 (2004).
151. Barrios, A. F. G. *et al.* Autoinducer 2 Controls Biofilm Formation in *Escherichia coli* through a Novel Motility Quorum-Sensing Regulator (MqsR, B3022). *J. Bacteriol.* **188**, 305 (2006).
152. Wang, X. *et al.* Antitoxin MqsA Helps Mediate the Bacterial General Stress Response. *Nat. Chem. Biol.* **7**, 359 (2011).
153. Brown, B. L. *et al.* Three Dimensional Structure of the MqsR:MqsA Complex: A Novel TA Pair Comprised of a Toxin Homologous to RelE and an Antitoxin with Unique Properties. *PLoS Pathog.* **5**, (2009).
154. Kim, Y. *et al. Escherichia coli* toxin/antitoxin pair MqsR/MqsA regulate toxin CspD. *Environ. Microbiol.* **12**, 1105 (2010).
155. Soo, V. W. C. & Wood, T. K. Antitoxin MqsA Represses Curli Formation Through the Master Biofilm Regulator CsgD. *Sci. Rep.* **3**, (2013).

156. Kim, Y., Wang, X., Ma, Q., Zhang, X.-S. & Wood, T. K. Toxin-Antitoxin Systems in *Escherichia coli* Influence Biofilm Formation through YjgK (TabA) and Fimbriae. *J. Bacteriol.* **191**, 1258 (2009).

157. Fang, F. C., Frawley, E. R., Tapscott, T. & Vazquez-Torres, A. Bacterial Stress Responses during Host Infection. *Cell Host Microbe* **20**, 133–143 (2016).

158. Poole, K. Bacterial stress responses as determinants of antimicrobial resistance. *J. Antimicrob. Chemother.* **67**, 2069–2089 (2012).

159. Kim, J. S. & Wood, T. K. Tolerant, Growing Cells from Nutrient Shifts Are Not Persister Cells. *MBio* **8**, e00354-17 (2017).

160. Wood, T. K., Knabel, S. J. & Kwan, B. W. Bacterial Persister Cell Formation and Dormancy. *Appl. Environ. Microbiol.* **79**, 7121 (2013).

161. Fisher, R. A., Gollan, B. & Helaine, S. Persistent bacterial infections and persister cells. *Nat. Rev. Microbiol.* **15**, 453–464 (2017).

162. Van Melderen, L. Toxin–antitoxin systems: why so many, what for? *Curr. Opin. Microbiol.* **13**, 781–785 (2010).

163. Maslowska, K. H., Makiela-Dzbenska, K. & Fijalkowska, I. J. The SOS system: A complex and tightly regulated response to DNA damage. *Environ. Mol. Mutagen.* **60**, 368–384 (2019).

164. Aldred, K. J., Kerns, R. J. & Osheroff, N. Mechanism of Quinolone Action and Resistance. *Biochemistry* **53**, 1565–1574 (2014).

165. Little, J. W. & Mount, D. W. The SOS Regulatory System of *Escherichia coli*. *Cell* **29**, 11 (1982).

166. Erill, I., Campoy, S. & Barbé, J. Aeons of distress: an evolutionary perspective on the bacterial SOS response. *FEMS Microbiol. Rev.* **31**, 637–656 (2007).

167. Henestrosa, A. R. F. de *et al.* Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol. Microbiol.* **35**, 1560–1572 (2000).

168. Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. & Hanawalt, P. C. Comparative Gene Expression Profiles Following UV Exposure in Wild-Type and SOS-Deficient *Escherichia coli*. *Genetics* **158**, 41–64 (2001).

169. Dörr, T., Vulić, M. & Lewis, K. Ciprofloxacin Causes Persister Formation by Inducing the TisB toxin in *Escherichia coli*. *PLOS Biol.* **8**, e1000317 (2010).

170. Prysak, M. H. *et al.* Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. *Mol. Microbiol.* **71**, 1071–1087 (2009).

171. Singletary, L. A. *et al.* An SOS-Regulated Type 2 Toxin-Antitoxin System. *J. Bacteriol.*

191, 7456 (2009).

172. Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).
173. Magnusson, L. U., Farewell, A. & Nyström, T. ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* **13**, 236–242 (2005).
174. Godfrey, H. P., Bugrysheva, J. V. & Cabello, F. C. The role of the stringent response in the pathogenesis of bacterial infections. *Trends Microbiol.* **10**, 349–351 (2002).
175. Pizarro-Cerdá, J. & Tedin, K. The bacterial signal molecule, ppGpp, regulates *Salmonella* virulence gene expression. *Mol. Microbiol.* **52**, 1827–1844 (2004).
176. Wu, J. & Xie, J. Magic spot: (p) ppGpp. *J. Cell. Physiol.* **220**, 297–302 (2009).
177. Hogg, T., Mechold, U., Malke, H., Cashel, M. & Hilgenfeld, R. Conformational Antagonism between Opposing Active Sites in a Bifunctional RelA/SpoT Homolog Modulates (p)ppGpp Metabolism during the Stringent Response. *Cell* **117**, 57–68 (2004).
178. Chang, D.-E., Smalley, D. J. & Conway, T. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. *Mol. Microbiol.* **45**, 289–306 (2002).
179. Sørensen, M. A., Jensen, K. F. & Pedersen, S. High Concentrations of ppGpp Decrease the RNA Chain Growth Rate: Implications for Protein Synthesis and Translational Fidelity During Amino Acid Starvation in *Escherichia coli*. *J. Mol. Biol.* **236**, 441–454 (1994).
180. Verstraeten, N. *et al.* Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Mol. Cell* **59**, 9–21 (2015).
181. Christensen-Dalsgaard, M., Jørgensen, M. G. & Gerdes, K. Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol. Microbiol.* **75**, 333–348 (2010).
182. Christensen, S. K., Mikkelsen, M., Pedersen, K. & Gerdes, K. RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14328 (2001).
183. Jørgensen, M. G., Pandey, D. P., Jaskolska, M. & Gerdes, K. HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J. Bacteriol.* **191**, 1191–1199 (2009).
184. Tosa, T. & Pizer, L. I. Biochemical bases for the antimetabolite action of L-serine hydroxamate. *J. Bacteriol.* **106**, 972–982 (1971).

185. Shan, Y. *et al.* ATP-Dependent persister formation in *Escherichia coli*. *MBio* **8**, (2017).

186. Svitils, A. L., Cashel, M. & Zyskindsll, J. W. Guanosine tetraphosphate inhibits protein synthesis *in vivo*. A possible protective mechanism for starvation stress in *Escherichia coli*. *J. Biol. Chem.* **268**, 2307–2311 (1993).

187. Tian, C. *et al.* Rapid Curtailing of the Stringent Response by Toxin-Antitoxin Module-Encoded mRNases. *J. Bacteriol.* **198**, 1918 (2016).

188. Lund, E. & Kjeldgaard, N. O. Metabolism of Guanosine Tetraphosphate in *Escherichia coli*. *Eur. J. Biochem.* **28**, 316–326 (1972).

189. Lazzarini, R. A., Cashel, M. & Gallant, J. On the Regulation of Guanosine Tetraphosphate Levels in Stringent and Relaxed Strains of *Escherichia coli*. *J. Biol. Chem.* **246**, 4381–4385 (1971).

190. Lagosky, P. A. & Chang, F. N. Influence of amino acid starvation on guanosine 5'-diphosphate 3'-diphosphate basal-level synthesis in *Escherichia coli*. *J. Bacteriol.* **144**, 499–508 (1980).

191. Traxler, M. F. *et al.* Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in *Escherichia coli*. *Mol. Microbiol.* **79**, 830–845 (2011).

192. Kwan, B. W. *et al.* The MqsR/MqsA toxin/antitoxin system protects *Escherichia coli* during bile acid stress. *Environ. Microbiol.* **17**, 3168–3181 (2015).

193. Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **27**, 3821–3835 (1999).

194. Figueira, R. *et al.* Adaptation to sustained nitrogen starvation by *Escherichia coli* requires the eukaryote-like serine/threonine kinase YeaG. *Sci. Rep.* **5**, 1–14 (2015).

195. Muranaka, L. S., Takita, M. A., Olivato, J. C., Kishi, L. T. & de Souza, A. A. Global expression profile of biofilm resistance to antimicrobial compounds in the plant-pathogenic bacterium *Xylella fastidiosa* reveals evidence of persister cells. *J. Bacteriol.* **194**, 4561–4569 (2012).

196. Merfa, M. V., Niza, B., Takita, M. A. & De Souza, A. A. The MqsRA Toxin-Antitoxin System from *Xylella fastidiosa* Plays a Key Role in Bacterial Fitness, Pathogenicity, and Persister Cell Formation. *Front. Microbiol.* **0**, 904 (2016).

197. Battesti, A., Majdalani, N. & Gottesman, S. The RpoS-Mediated General Stress Response in *Escherichia coli**. *Annu. Rev. Microbiol.* **65**, 189–213 (2011).

198. Fraikin, N., Goormaghtigh, F. & Van Melderen, L. Type II Toxin-Antitoxin Systems: Evolution and Revolutions. *J. Bacteriol.* **202**, (2020).

199. Fraikin, N., Rousseau, C. J., Goeders, N. & Van Melderen, L. Reassessing the role of the

type II MqsRA toxin-antitoxin system in stress response and biofilm formation: MqsA is transcriptionally uncoupled from *mqsR*. *MBio* **10**, (2019).

200. Tsilibaris, V., Maenhaut-Michel, G., Mine, N. & Van Melderen, L. What is the benefit to *Escherichia coli* of having multiple toxin-antitoxin systems in its genome? *J. Bacteriol.* **189**, 6101–6108 (2007).
201. Van Melderen, L. & Bast, M. S. De. Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities? *PLOS Genet.* **5**, e1000437 (2009).
202. Ramisetty, B. C. M. & Santhosh, R. S. Horizontal gene transfer of chromosomal Type II toxin–antitoxin systems of *Escherichia coli*. *FEMS Microbiol. Lett.* **363**, 238 (2016).
203. Wilbaux, M., Mine, N., Guérout, A.-M., Mazel, D. & Van Melderen, L. Functional interactions between coexisting toxin-antitoxin systems of the *ccd* family in *Escherichia coli* O157:H7. *J. Bacteriol.* **189**, 2712–2719 (2007).
204. Clément, J.-M., Wilde, C., Bachellier, S., Lambert, P. & Hofnung, M. IS1397 Is Active for Transposition into the Chromosome of *Escherichia coli* K-12 and Inserts Specifically into Palindromic Units of Bacterial Interspersed Mosaic Elements. *J. Bacteriol.* **181**, 6929 (1999).
205. Collis, C. M. & Hall, R. M. Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Mol. Microbiol.* **6**, 2875–2885 (1992).
206. Recchia, G. D. & Hall, R. M. Gene cassettes: a new class of mobile element. *Microbiology* **141**, 3015–3027 (1995).
207. Hall, R. M. & Stokes, H. W. Integrons: novel DNA elements which capture genes by site-specific recombination. *Genetica* **90**, 115–132 (1993).
208. Stokes, H. W. & Hall, R. M. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Mol. Microbiol.* **3**, 1669–1683 (1989).
209. Rowe-Magnus, D. A., Guérout, A.-M. & Mazel, D. Super-integrons. *Res. Microbiol.* **150**, 641–651 (1999).
210. Heidelberg, J. F. *et al.* DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477 (2000).
211. Chen, C.-Y. *et al.* Comparative Genome Analysis of *Vibrio vulnificus*, a Marine Pathogen. *Genome Res.* **13**, 2577 (2003).
212. Rowe-Magnus, D. A., Guerout, A.-M., Biskri, L., Bouige, P. & Mazel, D. Comparative Analysis of Superintegrons: Engineering Extensive Genetic Diversity in the Vibrionaceae. *Genome Res.* **13**, 428–442 (2003).
213. Christensen-Dalsgaard, M. & Gerdes, K. Two *higBA* loci in the *Vibrio cholerae*

superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Mol. Microbiol.* **62**, 397–411 (2006).

214. Huguet, K. T., Gonnet, M., Doublet, B. & Cloeckaert, A. A toxin antitoxin system promotes the maintenance of the IncA/C-mobilizable *Salmonella* Genomic Island 1. *Sci. Rep.* **6**, (2016).
215. Wozniak, R. A. F. & Waldor, M. K. A Toxin–Antitoxin System Promotes the Maintenance of an Integrative Conjugative Element. *PLoS Genet.* **5**, 1000439 (2009).
216. Yang, Q. E. & Walsh, T. R. Toxin–antitoxin systems and their role in disseminating and maintaining antimicrobial resistance. *FEMS Microbiol. Rev.* **41**, 343–353 (2017).
217. Santos Sierra, S., Giraldo, R. & Díaz Orejas, R. Functional interactions between *chpB* and *parD*, two homologous conditional killer systems found in the *Escherichia coli* chromosome and in plasmid R1. *FEMS Microbiol. Lett.* **168**, 51–58 (1998).
218. De Bast, M. S., Mine, N. & Van Melderen, L. Chromosomal toxin-antitoxin systems may act as antiaddiction modules. *J. Bacteriol.* **190**, 4603–4609 (2008).
219. Stewart, G. R., Robertson, B. D. & Young, D. B. Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* **1**, 97–105 (2003).
220. Grant, S. S. & Hung, D. T. Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. *Virulence* **4**, 273–283 (2013).
221. Brauner, A., Fridman, O., Gefen, O. & Balaban, N. Q. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat. Rev. Microbiol.* **14**, 320–330 (2016).
222. Levin-Reisman, I., Brauner, A., Ronin, I. & Balaban, N. Q. Epistasis between antibiotic tolerance, persistence, and resistance mutations. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 14734–14739 (2019).
223. Scholar, Eric M.; Pratt, W. B. *The Antimicrobial Drugs, Second Edition*. *Oxford University Press* (Oxford University Press, 2000).
224. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial Persistence as a Phenotypic Switch. *Science.* **305**, 1622–1625 (2004).
225. Shah, D. *et al.* Persisters: a distinct physiological state of *E. coli*. *BMC Microbiol.* **6**, 1–9 (2006).
226. Kussell, E., Kishony, R., Balaban, N. Q. & Leibler, S. Bacterial Persistence: A Model of Survival in Changing Environments. *Genetics* **169**, 1807 (2005).
227. Ventola, C. L. The Antibiotic Resistance Crisis: Part 1: Causes and Threats. *Pharm. Ther.* **40**, 277 (2015).
228. Windels, E. M. *et al.* Bacterial persistence promotes the evolution of antibiotic

resistance by increasing survival and mutation rates. *ISME J.* **13**, 1239–1251 (2019).

229. Jung, S.-H., Ryu, C.-M. & Kim, J.-S. Bacterial persistence: Fundamentals and clinical importance. *J. Microbiol.* **57**, 829–835 (2019).
230. Rotem, E. *et al.* Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12541–12546 (2010).
231. Vega, N. M., Allison, K. R., Khalil, A. S. & Collins, J. J. Signaling-mediated bacterial persister formation. *Nat. Chem. Biol.* **8**, 431–433 (2012).
232. Möker, N., Dean, C. R. & Tao, J. *Pseudomonas aeruginosa* increases formation of multidrug-tolerant persister cells in response to quorum-sensing signaling molecules. *J. Bacteriol.* **192**, 1946–1955 (2010).
233. Wu, Y., Vulić, M., Keren, I. & Lewis, K. Role of oxidative stress in persister tolerance. *Antimicrob. Agents Chemother.* **56**, 4922–4926 (2012).
234. Kotte, O., Volkmer, B., Radzikowski, J. L. & Heinemann, M. Phenotypic bistability in *Escherichia coli's* central carbon metabolism. *Mol. Syst. Biol.* **10**, 736 (2014).
235. Bernier, S. P. *et al.* Starvation, Together with the SOS Response, Mediates High Biofilm-Specific Tolerance to the Fluoroquinolone Ofloxacin. *PLOS Genet.* **9**, e1003144 (2013).
236. Harms, A., Maisonneuve, E. & Gerdes, K. Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science.* **354**, aaf4268 (2016).
237. Van Melderen, L. & Wood, T. K. Commentary: What Is the Link between Stringent Response, Endoribonuclease Encoding Type II Toxin-Antitoxin Systems and Persistence? *Front. Microbiol.* **8**, 191 (2017).
238. Ramisetty, B. C. M., Ghosh, D., Roy Chowdhury, M. & Santhosh, R. S. What Is the Link between Stringent Response, Endoribonuclease Encoding Type II Toxin–Antitoxin Systems and Persistence? *Front. Microbiol.* **7**, 1882 (2016).
239. Moyed, H. S. & Bertrand, K. P. *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J. Bacteriol.* **155**, 768–775 (1983).
240. Korch, S. B., Henderson, T. A. & Hill, T. M. Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol. Microbiol.* **50**, 1199–1213 (2003).
241. Germain, E., Castro-Roa, D., Zenkin, N. & Gerdes, K. Molecular Mechanism of Bacterial Persistence by HipA. *Mol. Cell* **52**, 248–254 (2013).
242. Kaspy, I. *et al.* HipA-mediated antibiotic persistence via phosphorylation of the

glutamyl-tRNA-synthetase. *Nat. Commun.* **4**, 1–7 (2013).

243. Semanjski, M. *et al.* The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Sci. Signal.* **11**, 5750 (2018).
244. Korch, S. B. & Hill, T. M. Ectopic overexpression of wild-type and mutant *hipA* genes in *Escherichia coli*: Effects on macromolecular synthesis and persister formation. *J. Bacteriol.* **188**, 3826–3836 (2006).
245. Keren, I., Shah, D., Spoering, A., Kaldalu, N. & Lewis, K. Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J. Bacteriol.* **186**, 8172–8180 (2004).
246. Vázquez-Laslop, N., Lee, H. & Neyfakh, A. A. Increased persistence in *Escherichia coli* caused by controlled expression of toxins or other unrelated proteins. *J. Bacteriol.* **188**, 3494–3497 (2006).
247. Harrison, J. J. *et al.* The chromosomal toxin gene *yafQ* is a determinant of multidrug tolerance for *Escherichia coli* growing in a biofilm. *Antimicrob. Agents Chemother.* **53**, 2253–2258 (2009).
248. Tripathi, A., Dewan, P. C., Siddique, S. A. & Varadarajan, R. MazF-induced Growth Inhibition and Persister Generation in *Escherichia coli**. *J. Biol. Chem.* **289**, 4191–4205 (2014).
249. Helaine, S. *et al.* Internalization of *Salmonella* by Macrophages Induces Formation of Nonreplicating Persisters. *Science.* **343**, 204–208 (2014).
250. Slattery, A., Victorsen, A. H., Brown, A., Hillman, K. & Phillips, G. J. Isolation of highly persistent mutants of *Salmonella enterica* serovar typhimurium reveals a new toxin-antitoxin module. *J. Bacteriol.* **195**, 647–657 (2013).
251. Torrey, H. L., Keren, I., Via, L. E., Lee, J. S. & Lewis, K. High Persister Mutants in *Mycobacterium tuberculosis*. *PLoS One* **11**, e0155127 (2016).
252. Keren, I., Minami, S., Rubin, E. & Lewis, K. Characterization and Transcriptome Analysis of *Mycobacterium tuberculosis* Persisters. *MBio* **2**, (2011).
253. Maisonneuve, E., Castro-Camargo, M. & Gerdes, K. Retraction: (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. *Cell* **172**, 1135 (2018).
254. Maisonneuve, E., Castro-Camargo, M. & Gerdes, K. (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. *Cell* **154**, 1140–1150 (2013).
255. Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G. & Gerdes, K. Bacterial persistence by RNA endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13206–13211

(2011).

256. Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G. & Gerdes, K. Retraction: Bacterial persistence by RNA endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2901 (2018).

257. Goormaghtigh, F. *et al.* Reassessing the role of type II toxin-antitoxin systems in formation of *Escherichia coli* type II persister cells. *MBio* **9**, (2018).

258. Harms, A., Fino, C., Sørensen, M. A., Semsey, S. & Gerdes, K. Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. *MBio* **8**, e01964-17 (2017).

259. Kaldalu, N., Hauryliuk, V. & Tenson, T. Persisters—as elusive as ever. *Appl. Microbiol. Biotechnol.* **100**, 6545–6553 (2016).

260. Walter, N. D. *et al.* Transcriptional Adaptation of Drug-tolerant *Mycobacterium tuberculosis* During Treatment of Human Tuberculosis. *J. Infect. Dis.* **212**, 990–998 (2015).

261. Blower, T. R., Salmond, G. P. C. & Luisi, B. F. Balancing at survival's edge: the structure and adaptive benefits of prokaryotic toxin–antitoxin partners. *Curr. Opin. Struct. Biol.* **21**, 109–118 (2011).

262. Magnuson, R. & Yarmolinsky, M. B. Corepression of the P1 Addiction Operon by Phd and Doc. *J. Bacteriol.* **180**, 6342 (1998).

263. Overgaard, M., Borch, J., Jørgensen, M. G. & Gerdes, K. Messenger RNA interferase RelE controls *relBE* transcription by conditional cooperativity. *Mol. Microbiol.* **69**, 841–857 (2008).

264. Garcia-Pino, A. *et al.* Allostery and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity. *Cell* **142**, 101–111 (2010).

265. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2014).

266. Garcia-Pino, A. *et al.* Doc of Prophage P1 Is Inhibited by Its Antitoxin Partner Phd through Fold Complementation*♦. *J. Biol. Chem.* **283**, 30821–30827 (2008).

267. Schureck, M. A. *et al.* Structure of the *Proteus vulgaris* HigB-(HigA)2-HigB Toxin-Antitoxin Complex*. *J. Biol. Chem.* **289**, 1060–1070 (2014).

268. Bøggild, A. *et al.* The Crystal Structure of the Intact *E. coli* RelBE Toxin-Antitoxin Complex Provides the Structural Basis for Conditional Cooperativity. *Structure* **20**, 1641–1648 (2012).

269. De Jonge, N. *et al.* Rejuvenation of CcdB-Poisoned Gyrase by an Intrinsically

Disordered Protein Domain. *Mol. Cell* **35**, 154–163 (2009).

270. Loris, R. *et al.* Crystal Structure of the Intrinsically Flexible Addiction Antidote MazE. *J. Biol. Chem.* **278**, 28252–28257 (2003).
271. Madl, T. *et al.* Structural Basis for Nucleic Acid and Toxin Recognition of the Bacterial Antitoxin CcdA. *J. Mol. Biol.* **364**, 170–185 (2006).
272. Oberer, M., Zangger, K., Gruber, K. & Keller, W. The solution structure of ParD, the antidote of the ParDE toxin–antitoxin module, provides the structural basis for DNA and toxin binding. *Protein Sci.* **16**, 1676–1688 (2007).
273. Loris, R. & Garcia-Pino, A. Disorder- and Dynamics-Based Regulatory Mechanisms in Toxin–Antitoxin Modules. *Chem. Rev.* **114**, 6933–6947 (2014).
274. Afif, H., Allali, N., Couturier, M. & Melderen, L. Van. The ratio between CcdA and CcdB modulates the transcriptional repression of the *ccd* poison–antidote system. *Mol. Microbiol.* **41**, 73–82 (2001).
275. Monti, M. C. *et al.* Interactions of Kid–Kis toxin–antitoxin complexes with the *parD* operator-promoter region of plasmid R1 are piloted by the Kis antitoxin and tuned by the stoichiometry of Kid–Kis oligomers. *Nucleic Acids Res.* **35**, 1737 (2007).
276. Li, G.-W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* **157**, 624–635 (2014).
277. Gelens, L., Hill, L., Vandervelde, A., Danckaert, J. & Loris, R. A General Model for Toxin-Antitoxin Module Dynamics Can Explain Persister Cell Formation in *E. coli*. *PLoS Comput. Biol.* **9**, e1003190 (2013).
278. Ruangprasert, A. *et al.* Mechanisms of Toxin Inhibition and Transcriptional Repression by *Escherichia coli* DinJ-YafQ. *J. Biol. Chem.* **289**, 20559–20569 (2014).
279. Brown, B. L., Lord, D. M., Grigorius, S., Peti, W. & Page, R. The *Escherichia coli* Toxin MqsR Destabilizes the Transcriptional Repression Complex Formed between the Antitoxin MqsA and the *mqsRA* Operon Promoter. *J. Biol. Chem.* **288**, 1286–1294 (2013).
280. Brantl, S. & Jahn, N. sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* **39**, 413–427 (2015).
281. Ingmer, H. & Brøndsted, L. Proteases in bacterial pathogenesis. *Res. Microbiol.* **160**, 704–710 (2009).
282. Guo, M. S. & Gross, C. A. Stress-Induced Remodeling of the Bacterial Proteome. *Curr. Biol.* **24**, R424–R434 (2014).
283. Pruteanu, M. & Baker, T. A. Proteolysis in the SOS response and metal homeostasis in

Escherichia coli. *Res. Microbiol.* **160**, 677–683 (2009).

284. Gottesman, S. Proteolysis in Bacterial Regulatory Circuits. *Annu. Rev. Cell Dev. Biol.* **19**, 565–587 (2003).
285. Sauer, R. T. *et al.* Sculpting the Proteome with AAA+ Proteases and Disassembly Machines. *Cell* **119**, 9–18 (2004).
286. Buts, L., Lah, J., Dao-Thi, M. H., Wyns, L. & Loris, R. Toxin–antitoxin modules as bacterial metabolic stress managers. *Trends Biochem. Sci.* **30**, 672–679 (2005).
287. Christensen, S. K., Pedersen, K., Hansen, F. G. & Gerdes, K. Toxin–antitoxin Loci as Stress-response-elements: ChpAK/MazF and ChpBK Cleave Translated RNAs and are Counteracted by tmRNA. *J. Mol. Biol.* **332**, 809–819 (2003).
288. Lobato-Márquez, D., Moreno-Córdoba, I., Figueroa, V., Díaz-Orejas, R. & García-del Portillo, F. Distinct type I and type II toxin-antitoxin modules control *Salmonella* lifestyle inside eukaryotic cells. *Sci. Rep.* **5**, 1–10 (2015).
289. Donegan, N. P., Thompson, E. T., Fu, Z. & Cheung, A. L. Proteolytic regulation of toxin-antitoxin systems by ClpPc in *Staphylococcus aureus*. *J. Bacteriol.* **192**, 1416–1422 (2010).
290. Lioy, V. S. *et al.* pSM19035-encoded ζ toxin induces stasis followed by death in a subpopulation of cells. *Microbiology* **152**, 2365–2379 (2006).
291. Hansen, S. *et al.* Regulation of the *Escherichia coli* HipBA Toxin-Antitoxin System by Proteolysis. *PLoS One* **7**, e39185 (2012).
292. Aizenman, E., Engelberg-Kulka, H. & Glaser, G. An *Escherichia coli* chromosomal 'addiction module' regulated by 3',5'-bispyrophosphate: A model for programmed bacterial cell death (stringent response/protein degradation/nutritional stress/programmed cell death). *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6059–6063 (1996).
293. Smith, A. S. G. & Rawlings, D. E. Efficiency of the pTF-FC2 *pas* Poison-antidote stability system in *Escherichia coli* is affected by the host strain, and antidote degradation requires the Lon protease. *J. Bacteriol.* **180**, 5458–5462 (1998).
294. Tsuchimoto, S., Nishimura, Y. & Ohtsubo, E. The stable maintenance system *pem* of plasmid R100: Degradation of PemI protein may allow PemK protein to inhibit cell growth. *J. Bacteriol.* **174**, 4205–4211 (1992).
295. Diago-Navarro, E., Hernández-Arriaga, A. M., Kubik, S., Konieczny, I. & Díaz-Orejas, R. Cleavage of the antitoxin of the *parD* toxin–antitoxin system is determined by the ClpAP protease and is modulated by the relative ratio of the toxin and the antitoxin. *Plasmid* **70**, 78–85 (2013).
296. Brzozowska, I. & Zielenkiewicz, U. Regulation of toxin–antitoxin systems by

proteolysis. *Plasmid* **70**, 33–41 (2013).

297. Lehnherr, H. & Yarmolinsky, M. B. Addiction protein Phd of plasmid prophage P1 is a substrate of the ClpXP serine protease of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3274–3277 (1995).

298. Grønlund, H. & Gerdes, K. Toxin-antitoxin systems homologous with *relBE* of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes. *J. Mol. Biol.* **285**, 1401–1415 (1999).

299. Ning, D., Ye, S., Liu, B. & Chang, J. The Proteolytic Activation of the *relNEs* (*ssr1114*/*slr0664*) Toxin–Antitoxin System by Both Proteases Lons and ClpP2s/Xs of *Synechocystis* sp. PCC 6803. *Curr. Microbiol.* **63**, (2011).

300. Koga, M., Otsuka, Y., Lemire, S. & Yonesaki, T. *Escherichia coli rnlA* and *rnlB* Compose a Novel Toxin–Antitoxin System. *Genetics* **187**, 123–130 (2011).

301. Winther, K. S. & Gerdes, K. Regulation of Enteric *vapBC* Transcription: Induction by VapC Toxin Dimer-Breaking. *Nucleic Acids Res.* **40**, 4347–4357 (2012).

302. Christensen, S. K. *et al.* Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: involvement of the *yefM-yoeB* toxin-antitoxin system. *Mol. Microbiol.* **51**, 1705–1717 (2004).

303. Ruiz-Echevarría, M. J., de la Cueva, G. & Díaz-Orejas, R. Translational coupling and limited degradation of a polycistronic messenger modulate differential gene expression in the *parD* stability system of plasmid R1. *Mol. Gen. Genet. MGG* **248**, 599–609 (1995).

304. Mutschler, H., Reinstein, J. & Meinhart, A. Assembly Dynamics and Stability of the Pneumococcal Epsilon Zeta Antitoxin Toxin (PezAT) System from *Streptococcus pneumoniae*. *J. Biol. Chem.* **285**, 21797–21806 (2010).

305. Overgaard, M., Borch, J. & Gerdes, K. RelB and RelE of *Escherichia coli* Form a Tight Complex That Represses Transcription via the Ribbon–Helix–Helix Motif in RelB. *J. Mol. Biol.* **394**, 183–196 (2009).

306. Gerdes, K. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J. Bacteriol.* **182**, 561–572 (2000).

307. Cambau, E. & Drancourt, M. Steps towards the discovery of *Mycobacterium tuberculosis* by Robert Koch, 1882. *Clin. Microbiol. Infect.* **20** 196–201 (2014).

308. Smith, I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* **16** 463–496 (2003).

309. Koch, R. Die Ätiologie der Tuberkulose. *Berliner Klin. Wochenschrift* **15**, 221–230 (1882).

310. Hershkovitz, I. *et al.* Detection and Molecular Characterization of 9000-Year-Old *Mycobacterium tuberculosis* from a Neolithic Settlement in the Eastern Mediterranean. *PLoS One* **3**, e3426 (2008).

311. Nicklisch, N. *et al.* Rib lesions in skeletons from early neolithic sites in Central Germany: On the trail of tuberculosis at the onset of agriculture. *Am. J. Phys. Anthropol.* **149**, 391–404 (2012).

312. World Health Organization. *Global Tuberculosis Report 2020*. (2020).

313. World Health Organization. The top 10 causes of death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (2020).

314. Pieters, J. *Mycobacterium tuberculosis* and the Macrophage: Maintaining a Balance. *Cell Host Microbe* **3**, 399–407 (2008).

315. Cambier, C. J., Falkow, S. & Ramakrishnan, L. Host Evasion and Exploitation Schemes of *Mycobacterium tuberculosis*. *Cell* **159**, 1497–1509 (2014).

316. Glickman, M. S. & Jacobs Jr., W. R. Microbial Pathogenesis Review of *Mycobacterium tuberculosis*: Dawn of a Discipline. *Cell* **104**, 477–485 (2001).

317. Ehrt, S. & Schnappinger, D. Mycobacterial survival strategies in the phagosome: defence against host stresses. *Cell. Microbiol.* **11**, 1170–1178 (2009).

318. Houben, D. *et al.* ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell. Microbiol.* **14**, 1287–1298 (2012).

319. Wel, N. van der *et al. M. tuberculosis* and *M. leprae* Translocate from the Phagolysosome to the Cytosol in Myeloid Cells. *Cell* **129**, 1287–1298 (2007).

320. Chai, Q., Zhang, Y. & Liu, C. H. *Mycobacterium tuberculosis*: An adaptable pathogen associated with multiple human diseases. *Front. Cell. Infect. Microbiol.* **8** 158 (2018).

321. Forrellad, M. A. *et al.* Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence* **4**, 3–66 (2012).

322. Galagan, J. E. *et al.* The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* **499**, 178–183 (2013).

323. Rohde, K. H., Veiga, D. F. T., Caldwell, S., Balázsi, G. & Russell, D. G. Linking the Transcriptional Profiles and the Physiological States of *Mycobacterium tuberculosis* during an Extended Intracellular Infection. *PLOS Pathog.* **8**, e1002769 (2012).

324. Voskuil, M. I. *et al.* Inhibition of Respiration by Nitric Oxide Induces a *Mycobacterium tuberculosis* Dormancy Program. *J. Exp. Med.* **198**, 705–713 (2003).

325. Rustad, T. R., Harrell, M. I., Liao, R. & Sherman, D. R. The Enduring Hypoxic Response of *Mycobacterium tuberculosis*. *PLoS One* **3**, e1502 (2008).

326. Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. & Duncan, K. Evaluation of a

nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* **43**, 717–731 (2002).

327. Gröschel, M. I., Sayes, F., Simeone, R., Majlessi, L. & Brosch, R. ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat. Rev. Microbiol.* **14**, 677–691 (2016).
328. World Health Organization. *The End TB Strategy: global strategy and targets for tuberculosis prevention, care and control after 2015*. (2014).
329. United Nations. *Global Sustainable Development Report 2016 Edition*. (2016).
330. Ginsberg, A. M. & Spigelman, M. Challenges in tuberculosis drug research and development. *Nat. Med.* **13**, 290–294 (2007).
331. World Health Organization. *Meeting report of the WHO expert consultation on the definition of extensively drug-resistant tuberculosis*. (2020).
332. Slayden, R. A., Dawson, C. C. & Cummings, J. E. Toxin–antitoxin systems and regulatory mechanisms in *Mycobacterium tuberculosis*. *Pathog. Dis.* **76**, fty039 (2018).
333. Albrethsen, J. *et al.* Proteomic Profiling of *Mycobacterium tuberculosis* Identifies Nutrient-starvation-responsive Toxin–antitoxin Systems. *Mol. Cell. Proteomics* **12**, 1180–1191 (2013).
334. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 (2003).
335. DeJesus, M. A. *et al.* Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome via Saturating Transposon Mutagenesis. *MBio* **8**, e02133-16 (2017).
336. Sberro, H. *et al.* Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning. *Mol. Cell* **50**, 136–148 (2013).
337. Guzman, L.-M., Belin, D., Carson, M. J. & Beckwith, J. Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
338. Lodge, J., Fear, J., Busby, S., Gunasekaran, P. & Kamini, N. R. Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol. Lett.* **95**, 271–276 (1992).
339. Aslanidis, C. & De Jong, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* **18**, 6069–6074 (1990).
340. Peränen, J., Rikkonen, M., Hyvönen, M. & Kääriäinen, L. T7 Vectors with a Modified T7*lac* Promoter for Expression of Proteins in *Escherichia coli*. *Anal. Biochem.* **236**, 371–373 (1996).

341. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

342. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28 (1990).

343. Im, H., Sambrook, J. & Russell, D. W. The Inoue Method for Preparation and Transformation of Competent *Escherichia. coli*: "Ultra Competent"; Cells. *BIO-PROTOCOL* **1**, (2011).

344. Schaefer, J., Jovanovic, G., Kotta-Loizou, I. & Buck, M. A data comparison between a traditional and the single-step β-galactosidase assay. *Data Br.* **8**, 350–352 (2016).

345. Schaefer, J., Jovanovic, G., Kotta-Loizou, I. & Buck, M. Single-step method for β-galactosidase assays in *Escherichia coli* using a 96-well microplate reader. *Anal. Biochem.* **503**, 56–57 (2016).

346. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).

347. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

348. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67** 235–242 (2011).

349. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64** 112–122 (2008).

350. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

351. Rodríguez, D. D. *et al.* Crystallographic *ab initio* protein structure solution below atomic resolution. *Nat. Methods* **6**, 651–653 (2009).

352. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Struct. Biol.* **62**, 1002–1011 (2006).

353. Vagin, A. A. *et al.* REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–2195 (2004).

354. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

355. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Struct. Biol.* **68**, 352–367 (2012).

356. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).

357. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta*

Crystallogr. D Biol. Crystallogr. **60**, 2126–2132 (2004).

358. Gore, S., Velankar, S. & Kleywegt, G. J. Implementing an X-ray validation pipeline for the Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 478–483 (2012).
359. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
360. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233** 123–138 (1993).
361. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
362. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
363. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
364. Li, F. *et al.* Crystal Structures of the *Bacillus stearothermophilus* CCA-Adding Enzyme and Its Complexes with ATP or CTP. *Cell* **111**, 815–824 (2002).
365. Cole, S. T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
366. Logsdon, M. M. & Aldridge, B. B. Stable regulation of cell cycle events in mycobacteria: Insights from inherently heterogeneous bacterial populations. *Front. Microbiol.* **9**, 00514 (2018).
367. Sawyer, E. B., Grabowska, A. D. & Cortes, T. Translational regulation in mycobacteria and its implications for pathogenicity. *Nucleic Acids Res.* **46**, 6950–6961 (2018).
368. Gupta, A. Killing activity and rescue function of genome-wide toxin–antitoxin loci of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **290**, 45–53 (2009).
369. Zhu, L. *et al.* Characterization of mRNA Interferases from *Mycobacterium tuberculosis**. *J. Biol. Chem.* **281**, 18638–18643 (2006).
370. Singh, R., Barry, C. E. & Boshoff, H. I. M. The Three RelE Homologs of *Mycobacterium tuberculosis* Have Individual, Drug-Specific Effects on Bacterial Antibiotic Tolerance. *J. Bacteriol.* **192**, 1279–1291 (2010).
371. Zhu, L., Sharp, J. D., Kobayashi, H., Woychik, N. A. & Inouye, M. Noncognate *Mycobacterium tuberculosis* Toxin-Antitoxins Can Physically and Functionally Interact*. *J. Biol. Chem.* **285**, 39732–39738 (2010).
372. Carroll, P., Brown, A. C., Hartridge, A. R. & Parish, T. Expression of *Mycobacterium*

tuberculosis Rv1991c using an arabinose-inducible promoter demonstrates its role as a toxin. *FEMS Microbiol. Lett.* **274**, 73–82 (2007).

373. Kolodkin-Gal, I., Verdiger, R., Shlosberg-Fedida, A. & Engelberg-Kulka, H. A Differential Effect of *E. coli* Toxin-Antitoxin Systems on Cell Death in Liquid Media and Biofilm Formation. *PLoS One* **4**, e6785 (2009).
374. Jin, G., Pavelka, M. S. & Butler, J. S. Structure-function analysis of VapB4 antitoxin identifies critical features of a minimal VapC4 toxin-binding module. *J. Bacteriol.* **197**, 1197–207 (2015).
375. Norton, J. P. & Mulvey, M. A. Toxin-Antitoxin Systems Are Important for Niche-Specific Colonization and Stress Resistance of Uropathogenic *Escherichia coli*. *PLoS Pathog.* **8**, e1002954 (2012).
376. Schoch, C. L. *et al.* NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)* baaa062 (2020).
377. Frampton, R., Aggio, R. B. M., Villas-Bôas, S. G., Arcus, V. L. & Cook, G. M. Toxin-antitoxin systems of *Mycobacterium smegmatis* are essential for cell survival. *J. Biol. Chem.* **287**, 5340–56 (2012).
378. Janowski, R., Panjikar, S., Eddine, A. N., Kaufmann, S. H. E. & Weiss, M. S. Structural analysis reveals DNA binding properties of Rv2827c, a hypothetical protein from *Mycobacterium tuberculosis*. *J. Struct. Funct. Genomics* **10**, 137–150 (2009).
379. Smale, S. T. Beta-galactosidase assay. *Cold Spring Harb. Protoc.* (2010).
380. Beck, I. N., Usher, B., Hampton, H. G., Fineran, P. C. & Blower, T. R. Antitoxin autoregulation of *M. tuberculosis* toxin-antitoxin expression through negative cooperativity arising from multiple inverted repeat sequences. *Biochem. J.* **477**, 2401–2419 (2020).
381. Cortes, T. *et al.* Genome-wide Mapping of Transcriptional Start Sites Defines an Extensive Leaderless Transcriptome in *Mycobacterium tuberculosis*. *Cell Rep.* **5**, 1121–1131 (2013).
382. Shell, S. S. *et al.* Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. *PLOS Genet.* e1005641 (2015).
383. Agarwal, N. & Tyagi, A. K. Mycobacterial transcriptional signals: requirements for recognition by RNA polymerase and optimal transcriptional activity. *Nucleic Acids Res.* **34**, 4245–4257 (2006).
384. Rustad, T. R. *et al.* Mapping and manipulating the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biol.* **15**, 502 (2014).

385. Sachdeva, P., Misra, R., Tyagi, A. K. & Singh, Y. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J.* **277**, 605–626 (2010).

386. Akarsu, H. *et al.* TASmania: A bacterial Toxin-Antitoxin Systems database. *PLOS Comput. Biol.* **15**, e1006946 (2019).

387. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **27**, 129–134 (2018).

388. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

389. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

390. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

391. Alber, T. Signaling mechanisms of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinases. *Curr. Opin. Struct. Biol.* **19**, 650–657 (2009).

392. Molle, V. & Kremer, L. Division and cell envelope regulation by Ser/Thr phosphorylation: *Mycobacterium* shows the way. *Mol. Microbiol.* **75**, 1064–1077 (2010).

393. Domian, I. J., Quon, K. C. & Shapiro, L. Cell Type-Specific Phosphorylation and Proteolysis of a Transcriptional Regulator Controls the G1-to-S Transition in a Bacterial Cell Cycle. *Cell* **90**, 415–424 (1997).

394. Kobir, A. *et al.* Protein phosphorylation in bacterial signal transduction. *Biochim. Biophys. Acta* **1810**, 989–994 (2011).

395. Dworkin, J. Ser/Thr phosphorylation as a regulatory mechanism in bacteria. *Curr. Opin. Microbiol.* **24**, 47–52 (2015).

396. Trentini, D. B. *et al.* Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature* **539**, 48–53 (2016).

397. Hoch, J. A. Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* **3**, 165–170 (2000).

398. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-Component Signal Transduction. *Annu. Rev. Biochem.* **69**, 183–215 (2003).

399. Bonne Køhler, J. *et al.* Importance of protein Ser/Thr/Tyr phosphorylation for bacterial pathogenesis. *FEBS Lett.* **594**, 2339–2369 (2020).

400. Groisman, E. A. The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol.* **183**, 1835–1842 (2001).

401. Scheeff, E. D. & Bourne, P. E. Structural Evolution of the Protein Kinase–Like Superfamily. *PLOS Comput. Biol.* **1**, e49 (2005).

402. Durocher, D. Bacterial signal transduction: a FHAscinating glimpse at the origins of phospho-dependent signal transduction. *Trends Microbiol.* **11**, 67–68 (2003).

403. Nishi, H., Hashimoto, K. & Panchenko, A. R. Phosphorylation in protein-protein binding: effect on stability and function. *Structure* **19**, 1807–1815 (2011).

404. Zhao, Y. *et al.* Crystal Structure Confirmation of JHP933 as a Nucleotidyltransferase Superfamily Protein from *Helicobacter pylori* Strain J99. *PLoS One* **9**, e104609 (2014).

405. Yoon, J. Y. *et al.* Crystal structure of JHP933 from *Helicobacter pylori* J99 shows two-domain architecture with a DUF1814 family nucleotidyltransferase domain and a helical bundle domain. *Proteins Struct. Funct. Bioinforma.* **82**, 2275–2281 (2014).

406. Ye, X. *et al.* Identification of residues involved in nucleotidyltransferase activity of JHP933 from *helicobacter pylori* by site-directed mutagenesis. *Arch. Biol. Sci.* **68**, 439–444 (2016).

407. Lunde, B. M., Magler, I. & Meinhart, A. Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res.* **40**, 9815–9824 (2012).

408. Hampton, H. G. *et al.* Functional genomics reveals the toxin–antitoxin repertoire and AbiE activity in *Serratia*. *Microb. Genomics* **6**, 1–15 (2020).

409. Kurata, T. *et al.* RelA-SpoT Homolog toxins pyrophosphorylate the CCA end of tRNA to inhibit protein synthesis. *Mol. Cell* **81**, 3160-3170.e9 (2021).

410. Songailiene, I. *et al.* HEPN-MNT Toxin-Antitoxin System: The HEPN Ribonuclease Is Neutralized by OligoAMPylation. *Mol. Cell* **80**, 955-970.e7 (2020).

411. Barik, S., Sureka, K., Mukherjee, P., Basu, J. & Kundu, M. RseA, the SigE specific anti-sigma factor of *Mycobacterium tuberculosis*, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol. Microbiol.* **75**, 592–606 (2010).

412. Prisic, S. & Husson, R. N. *Mycobacterium tuberculosis* Serine/Threonine Protein Kinases . *Microbiol. Spectr.* **2**, (2014).

413. Iswahyudi *et al.* Mycobacterial phosphatase PstP regulates global serine threonine phosphorylation and cell division. *Sci. Rep.* **9**, 1–11 (2019).

414. Sharma, A. K. *et al.* Serine/threonine protein phosphatase *pstp* of *mycobacterium tuberculosis* is necessary for accurate cell division and survival of pathogen. *J. Biol. Chem.* **291**, 24215–24230 (2016).

415. Missiakas, D. & Raina, S. Signal transduction pathways in response to protein misfolding in the extracytoplasmic compartments of *E.coli*: role of two new

phosphoprotein phosphatases PrpA and PrpB. *EMBO J.* **16**, 1670–1685 (1997).

416. Rajagopalan, K. & Dworkin, J. Identification and biochemical characterization of a novel protein phosphatase 2C-like Ser/Thr phosphatase in *Escherichia coli*. *J. Bacteriol.* **200**, (2018).
417. Walling, L. R. & Butler, J. S. Toxins targeting tRNAs: translation inhibition by bacterial toxin-antitoxin systems. *Wiley Interdiscip. Rev. RNA* **10**, e1506 (2019).
418. Jurėnas, D. & Van Melderen, L. The Variety in the Common Theme of Translation Inhibition by Type II Toxin–Antitoxin Systems. *Front. Genet.* **11**, 262 (2020).
419. Gupta, A., Venkataraman, B., Vasudevan, M. & Gopinath Bankar, K. Co-expression network analysis of toxin-antitoxin loci in *Mycobacterium tuberculosis* reveals key modulators of cellular stress. *Sci. Rep.* **7**, 1–14 (2017).
420. Walling, L. R. & Butler, J. S. Homologous VapC toxins inhibit translation and cell growth by sequence-specific cleavage of $tRNA^{fMet}$. *J. Bacteriol.* **200**, e00582-17 (2018).
421. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–D189 (2016).
422. Samhita, L., Nanjundiah, V. & Varshneya, U. How many initiator tRNA genes does *Escherichia coli* need? *J. Bacteriol.* **196**, 2607–2615 (2014).
423. Shiloh, M. U. & DiGiuseppe Champion, P. A. To catch a killer. What can mycobacterial models teach us about *Mycobacterium tuberculosis* pathogenesis? *Curr. Opin. Microbiol.* **13**, 86–92 (2010).
424. Reyrat, J. M. & Kahn, D. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends Microbiol.* **9**, 472–473 (2001).
425. Barry, C. E. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends Microbiol.* **9**, 473–474 (2001).
426. Tyagi, J. S. & Sharma, D. *Mycobacterium smegmatis* and tuberculosis. *Trends Microbiol.* **10**, 68–69 (2002).
427. Srivastava, A. *et al.* Reconstitution of Protein Translation of *Mycobacterium* Reveals Functional Conservation and Divergence with the Gram-Negative Bacterium *Escherichia coli*. *PLoS One* **11**, e0162020 (2016).
428. Bordes, P. & Genevaux, P. Control of Toxin-Antitoxin Systems by Proteases in *Mycobacterium Tuberculosis*. *Front. Mol. Biosci.* **8**, 458 (2021).

Appendix

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

PORTLAND PRESS

Research Article

Antitoxin autoregulation of *M. tuberculosis* toxin-antitoxin expression through negative cooperativity arising from multiple inverted repeat sequences

Izaak N. Beck[1], Ben Usher[1], Hannah G. Hampton[2], Peter C. Fineran[2] and Tim R. Blower[1]

[1]Department of Biosciences, Durham University, Stockton Road, Durham DH1 3LE, U.K.; [2]Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin 9054, New Zealand

Correspondence: Tim R. Blower (timothy.blower@durham.ac.uk)

Toxin-antitoxin systems play key roles in bacterial adaptation, including protection from antibiotic assault and infection by bacteriophages. The type IV toxin-antitoxin system AbiE encodes a DUF1814 nucleotidyltransferase-like toxin, and a two-domain antitoxin. In *Streptococcus agalactiae*, the antitoxin AbiEi negatively autoregulates *abiE* expression through positively co-operative binding to inverted repeats within the promoter. The human pathogen *Mycobacterium tuberculosis* encodes four DUF1814 putative toxins, two of which have antitoxins homologous to AbiEi. One such *M. tuberculosis* antitoxin, named Rv2827c, is required for growth and whilst the structure has previously been solved, the mode of regulation is unknown. To complete the gaps in our understanding, we first solved the structure of *S. agalactiae* AbiEi to 1.83 Å resolution for comparison with *M. tuberculosis* Rv2827c. AbiEi contains an N-terminal DNA binding domain and C-terminal antitoxicity domain, with bilateral faces of opposing charge. The overall AbiEi fold is similar to Rv2827c, though smaller, and with a 65° difference in C-terminal domain orientation. We further demonstrate that, like AbiEi, Rv2827c can autoregulate toxin-antitoxin operon expression. In contrast with AbiEi, the $P_{rv2827c}$ promoter contains two sets of inverted repeats, which bind Rv2827c with differing affinities depending on the sequence consensus. Surprisingly, Rv2827c bound with negative co-operativity to the full $P_{rv2827c}$ promoter, demonstrating an unexpectedly complex form of transcriptional regulation.

Introduction

Toxin-antitoxin (TA) systems are encoded by genetic loci that are widely distributed throughout prokaryotic genomes. They can play pivotal roles in bacterial physiology and in managing stress responses, helping bacteria to survive nutrient limitation, immune system attack, antibiotic treatment and predation by bacteriophages [1–5]. TA systems are commonly found on mobile genetic elements, contributing to the stability of plasmids, superintegrons, cryptic prophages and conjugative transposons [6–8]. The majority of TA systems encode two components, a toxic protein that generally targets essential cellular processes, and an antagonistic antitoxin [4]. This antitoxin negates toxin activity when cells are growing in favorable conditions. Under stressful conditions, the antitoxin is preferentially degraded and the toxin is released, thereby reducing growth rate as a means to survive with minimal metabolic burden until favorable conditions return [9,10]. Activation of the toxin following bacteriophage infection can also lead to the removal of the infectious bacteriophage particle from the environment, thereby providing a population level protection from viruses referred to as abortive infection (Abi) [11,12].

TA systems have been divided into six types according to the nature of the toxin and antitoxin (whether they are RNA or protein), and the mechanism of toxin antagonism [4]. Type IV systems differ from all others in that the antitoxin and toxin do not directly interact, instead, the antitoxin

Received: 6 May 2020
Revised: 6 June 2020
Accepted: 10 June 2020

Accepted Manuscript online: 10 June 2020
Version of Record published: 26 June 2020

Biochemical Journal (2020) 477 2401–2419
https://doi.org/10.1042/BCJ20200368

antagonizes the activity of the toxin [13–15]. There are multiple examples wherein TA systems provide a phage-resistant Abi phenotype, although not all identified Abi systems act as bona fide TA systems [5,15–20]. A recently characterized Abi system, AbiE from *Streptococcus agalactiae* V/R 2603, has been shown to act as a type IV TA system [15]. AbiE encodes a DUF1814-family toxin (AbiEii), and a COG5340-family antitoxin (AbiEi) (Figure 1A) [15]. The *S. agalactiae* AbiE COG5340 antitoxin will herein be referred to as AbiEi. AbiEii is a putative nucleotidyltransferase (NTase) that specifically binds GTP [15]. This DUF1814 family is widespread, present in over 5500 bacterial, archaeal and fungal genomes, though not all examples are genetically linked to putative antitoxins.

TA systems are remarkably abundant in *Mycobacterium tuberculosis*, which encodes more than 80 examples, and these are thought to have contributed to *M. tuberculosis* having become the most successful human pathogen [21–24]. *M. tuberculosis* H37Rv encodes four DUF1814-family NTase-like putative toxins, namely Rv0078A, Rv0836c, Rv1045 and Rv2826c (Figure 1A). Akin to AbiEii from *S. agalactiae*, both Rv1045 and Rv2826c have a cognate COG5340-family antitoxin (Figure 1A). Transposon mutagenesis studies have identified the cognate antitoxins of these systems (Rv1044 and Rv2827c) as essential for laboratory growth [25,26], suggesting that Rv1045 and Rv2826c toxins are functional in *M. tuberculosis*. The *M. tuberculosis* COG5340 proteins will herein be referred to by their respective 'Rv' identifiers, Rv1044 and Rv2827c. Characterizing and understanding the regulation of these loci is of interest for developing new therapies against the pathogen.

Autoregulation of TA system expression is a hallmark of type II TA systems and can be either positive or negative [27,28]. The antitoxin AbiEi from *S. agalactiae* has been biochemically characterized [15,29] and functions as both an antitoxin and a transcriptional repressor. That is, AbiEi negatively autoregulates *abiE* expression. Here, the gene product suppresses its own production, through positively co-operative binding of two AbiEi monomers to inverted repeats in the promoter region. Full length AbiEi is required for negative autoregulation and induced bending of the promoter DNA. We previously proposed that this bending was facilitated by the two AbiEi monomers interacting via their C-terminal domains (CTDs) [29]. In contrast with type II autoregulation, for which conditional co-operativity is observed, co-expression of the cognate toxin AbiEii does not enhance transcriptional repression [15]. We therefore sought to determine the similarities in the structure and function of AbiEi and Rv2827c. While the structure of the *M. tuberculosis* putative antitoxin Rv2827c has been solved as part of a structural genomics initiative [30], its biological function was not explored and it has not been biochemically characterized. We present the solved structure of *S. agalactiae* AbiEi, demonstrating structural homology between the COG5340 antitoxins, and biochemically characterize the molecular interactions underpinning transcriptional repression by Rv2827c. Interestingly, this is a more complex autoregulatory system than previously seen for AbiEi [29].

Materials and methods

Bacterial strains and culture conditions

E. coli DH5α (Invitrogen), BL21 (DE3) (Invitrogen) and ER2566 (New England Biolabs) were routinely grown at 37°C in Luria-Broth (LB), M9 minimal (M9M), or 2× YT media supplemented when necessary with ampicillin (Ap, 50 μg/ml), spectinomycin (Sp, 100 μg/ml), tetracycline (Tc, 10 μg/ml), isopropyl-β-D-thiogalactopyranoside (IPTG, 1 mM), L-arabinose (L-ara, 0.1% w/v) or D-glucose (glu, 0.2% w/v). Bacterial cell density was measured using a WPA Biowave C08000 at 600 nm (OD_{600}).

DNA isolation and manipulation

All oligonucleotides used in this study were obtained from Integrated DNA Technologies (Supplementary Table S1). Plasmid and PCR-amplified DNAs were purified using Monarch kits (NEB). Digests, ligations, transformations and agarose gel electrophoresis steps were performed by standard techniques. All constructed plasmids (Supplementary Table S2) were confirmed by sequencing using an ABI 3730 DNA sequencer and 4Peaks.

Protein expression constructs were made by Ligation Independent Cloning (LIC) [31]. Target genes were cloned into plasmid pSAT1-LIC, which generates N-terminal His_6-SUMO fusions with the target ORF. Primers TRB1048/TRB1049 were used to amplify *abiEi* from pRLD30, for LIC insertion into pSAT1-LIC, producing pTRB525. Primers TRB1022/TRB1023 were used to amplify *rv2827c* from pPF658, also for LIC insertion into pSAT1-LIC, producing pTRB493. Primers TRB1018/TRB1019 were used to amplify *rv1044* from *M. tuberculosis* H37Rv genomic DNA (ATCC), again for LIC insertion into pSAT1-LIC, producing pTRB491.

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

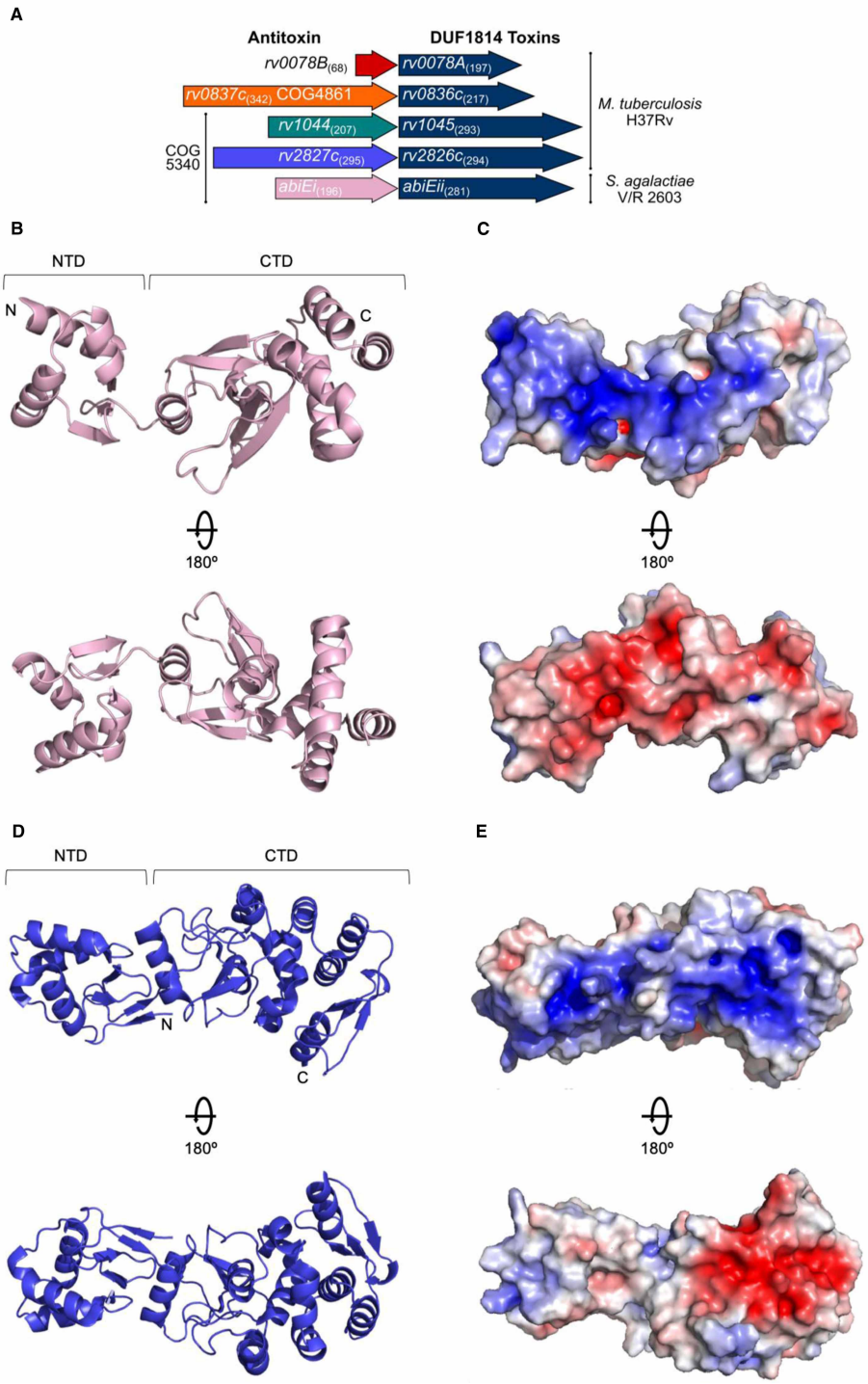


Figure 1. Antitoxin AbiEi is a two-domain protein with bilateral opposingly charged faces.

(**A**) Scaled representation of the four *M. tuberculosis* TA systems containing NTase-like toxin genes and AbiE from *S. agalactiae*. Numbers in parentheses indicate amino acid length. All five toxins are DUF1814 proteins; Rv1044, Rv2827c and AbiEi are COG5340-containing antitoxins. Putative antitoxin Rv0837c is a COG4861 protein and the significantly shorter putative antitoxin Rv0078B is unclassified. The four *M. tuberculosis* systems were re-named as shown. (**B**) AbiEi antitoxin structure shown in pink cartoon representation, in two views rotated 180°. (**C**) Electrostatic potential of AbiEi, posed as per (**B**), with electropositive charge in blue and electronegative charge in red. (**D**) Previously solved Rv2827c structure shown in blue cartoon representation, in two views rotated 180° (PDB: 1ZEL). (**E**) Electrostatic potential of Rv2827c, posed as per (**D**), colored as per (**C**).

Biochemical Journal (2020) 477 2401–2419
https://doi.org/10.1042/BCJ20200368

For promoter activity assays, regions upstream of *abiEi*, *rv2827c* and *rv1044* were cloned into pRW50 [32]. The 99 bp region upstream of *abiEi* was amplified from pPF680 using primers TRB1072/TRB1047, then digested with EcoRI/HindIII and ligated into pRW50 cut with the same enzymes, producing pTRB486. The 500 bp regions upstream of *rv2827c* and *rv1044* were amplified from H37Rv genomic DNA, using primers TRB1042/TRB1043 and TRB1040/TRB1041, respectively. The amplicons were digested with EcoRI/HindIII and ligated into pRW50 cut with the same enzymes, producing pTRB484 and pTRB483, respectively. Antitoxin genes *abiEi, rv2827c and rv1044* were cloned into pTA100, a pQE-80 derivative [5]. *S. agalactiae abiEi* was amplified from pRLD30 using primers TRB1052/TRB1053, then digested with EcoRI/HindIII and ligated into pTA100 cut with the same enzymes, producing pTRB481. *M. tuberculosis rv2827c* and *rv1044* were amplified from H37Rv genomic DNA, using primers PF1334/PF1335 and PF1330/PF1331, respectively. The amplicons were digested with NdeI/SpeI and ligated into pTA100 cut with the same enzymes, producing pPF658 and pPF658, respectively.

Protein expression and purification

To express AbiEi, Rv2827c and Rv1044 for crystallization and/or biochemistry, *E. coli* ER2566 (for native protein) or BL21 (DE3) (for labeled protein) were transformed with pTRB525, pTRB493 or pTRB491, respectively. For native protein, overnight cultures were re-seeded 1 : 100 into 2 L flasks containing 1 L 2× YT. Cells were grown at 150 rpm in 37°C until an OD_{600} of 0.6–0.8 was reached, whereupon expression was induced by the addition of IPTG (1 mM). Cells were left to grow for 16 h at 17°C, shaking at 150 rpm.

For incorporation of selenomethionine into AbiEi, the SeMet kit (Molecular Dimensions) was used. Starter cultures of BL21 (DE3) pTRB525, starter cultures were grown for 8 hours in LB at 37 °C with 200 rpm shaking. This culture was used to inoculate (1 : 500) a 50 ml overnight of Molecular Dimensions Selenomethionine Base medium supplemented with Molecular Dimensions Nutrient Mix. This overnight was then used to inoculate (1 : 100) 1 L of the same Base Medium with Nutrient Mix and cells were grown at 37°C with 180 rpm shaking. At OD_{600} 0.8, cells were pelleted by centrifugation at 4200×***g***, resuspended in fresh Base Medium with Nutrient Mix (Molecular Dimensions) and supplemented with an amino acid mix to promote feedback inhibition of methionine synthesis (0.1 mg/ml L-lysine hydrate, 0.1 mg/ml L-threonine, 0.1 mg/ml L-phenylalanine, 0.05 mg/ml L-leucine, 0.05 mg/ml L-isoleucine, 0.05 mg/ml L-valine). Cells were grown for a further 30 min at 37°C with shaking at 180 rpm before the addition of 250× SelenoMethionine Solution (Molecular Dimensions) to a final concentration of 40 μg/ml. Cells were grown for a further 15 min at 37°C with shaking at 180 rpm before antitoxin expression was induced with IPTG (1 mM), and samples were left to grow overnight at 175 rpm in 18°C.

For native protein purification, bacteria were harvested by centrifugation at 4200×***g*** and the pellets were resuspended in buffer A500 (20 mM Tris–HCl pH 7.9, 500 mM NaCl, 5 mM imidazole and 10% glycerol). Cells were lysed by sonication at 40 kpsi, then centrifuged (45 000×***g***, 4°C). The clarified lysate was next passed over a HisTrap HP column (GE Healthcare), washed for ten column volumes with A500, followed by 10 column volumes of buffer A100 (20 mM Tris–HCl pH 7.9, 100 mM NaCl, 5 mM imidazole and 10% w/v glycerol), then eluted directly onto a HiTrap Q HP column (GE Healthcare) with buffer B100 (20 mM Tris–HCl pH 7.9, 100 mM NaCl, 250 mM imidazole and 10% w/v glycerol). The Q HP column was transferred to an Äkta Pure (GE Healthcare), washed with three column volumes of A100, then proteins were eluted using a gradient from 100% A100 to 100% buffer C1000 (20 mM Tris–HCl pH 7.9, 1000 mM NaCl and 10% w/v glycerol). Fractions containing the protein peak were analyzed by SDS–PAGE, pooled and incubated overnight at 4°C with hSENP2 SUMO protease (SENP) to cleave the His_6-SUMO tag from the target protein. The following day, the samples were passed through a second HisTrap HP column and the flow-through fractions containing untagged target protein were collected. The same procedure was used for labeled protein, except seleno-AbiEi precipitated on column at A100, so contaminants were removed with B100, and remaining folded seleno-AbiEi was eluted with B500, followed by SENP cleavage and a second HisTrap column purification. Proteins were dialyzed overnight at 4°C into buffer X (20 mM Tris–HCl pH 7.9, 200 mM NaCl and 2.5 mM DTT). Crystallization samples were concentrated, quantified and stored on ice, then either used immediately or flash-frozen in liquid N_2 for storage at −80°C.

Protein crystallization

Native and selenomethionine-derivatized AbiEi were concentrated to 12 mg/ml in buffer X (see above). Initial native AbiEi crystallization screens were performed using commercial screens (Molecular Dimensions) set with

an Innovadyne Screenmaker robot, making 200 : 100 nl and 100 : 100 nl protein:condition sitting drops at 18°C. After initial screening and optimization, native AbiEi formed thick needles in 0.02 M Sodium/Potassium phosphate, 0.1 M Bis-Tris Propane pH 6.5, 20% PEG 3350. Selenomethionine-derivatized AbiEi crystals grew in 0.2 M Sodium acetate trihydrate, 0.1 M Bis-Tris Propane pH 6.5, 15% PEG 3350. To harvest, 20 μl of condition reservoir was added to 20 μl of glycerol and mixed quickly by vortexing; an equal volume of this mixture was then added to the drop volume. After addition of cryo buffer, crystals were immediately extracted using a nylon loop and flash-frozen in liquid N_2.

X-ray data collection and structure determination

Diffraction data were collected at Diamond Light Source on beamline I03 (AbiEi native and AbiEi selenomethionine-derivatized) (Table 1). Single, 360°, datasets were collected from three native AbiEi crystals and merged using iSpyB (Diamond Light Source). Two, 360°, datasets from AbiEi selenomethionine-derivatized crystals measured at the selenium peak (0.9793 Å) were also merged using iSpyB. An additional AbiEi selenomethionine-derivatized dataset was collected at selenium high remote (0.9641 Å) wavelength. Diffraction data were processed with XDS [33,34], and then AIMLESS from CCP4 [35] was used to corroborate the space-groups (Table 1). The crystal structure of AbiEi was solved by MAD, by providing the SHELX suite [36] in CCP4 with the native, peak and high remote datasets. The solved starting model for AbiEi was built in REFMAC [37] and BUCCANEER [38]. The model was then iteratively refined and built using PHENIX [39] and COOT [40], respectively. The quality of the final model was assessed using COOT and the wwPDB validation server [41]. Structural figures were generated using PyMol (Schrödinger). Structural alignments were performed using PROMALS3D [42].

Electrophoretic mobility shift assays

Conservation of IR sequences was determined using MView [43] and WebLogo [44]. Promoter region probes were amplified from synthesized templates (Supplementary Table S1). Each template was made with a common downstream region, matching the initial part of *lacZ* from pRW50. For each probe template, a unique upstream forward primer was used in combination with a common reverse primer, which was either untagged or had been conjugated to a fluorescein tag for visualization (Supplementary Table S1). The probes contained either the native promoter regions, or combinations of WT IR sequences and mutant sequences of polyC.

Proteins were diluted to appropriate concentrations using diluent buffers matching their storage buffer constitution. Each binding reaction contained 2 μl of 5× EMSA binding buffer (750 mM KCl, 50 mM Tris–HCl pH 8.0, 2.5 mM EDTA pH 8.0, 0.5% Triton X-100, 1 mM DTT, 55% glycerol), 250 fmoles of fluorescently labeled probe, 0.1 μl BSA (10 mg/ml), 1 μl poly(d[IC]) (1 mg/ml), 1 μl of diluted protein or buffer control and water to a final volume of 10 μl. Native 0.5× TBE polyacrylamide gels (at either 7% or 5% acrylamide, as required) were pre-run at 150 V and 4°C for 2 h. Binding reactions were titrated at protein concentrations from zero to an appropriate upper limit, and incubated at 20°C for 30 min. Non-specific binding controls used an additional excess of 2.5 μM TRB1108 template DNA amplified with TRB1109 as forward primer, and non-labeled reverse primer. Specific binding controls used additional excess of 2.5 μM unlabeled specific probe DNA. Samples were then separated by native polyacrylamide gel electrophoresis at 200 V and 4°C for 45 min.

Native polyacrylamide gels were then visualized using the Amersham Biosciences Typhoon 9400 on variable mode image in fluorescence mode, emission filter 526 SP. Sensitivity was set to normal. Band intensities were calculated using the grid scan feature and triplicate data processed in Prism (GraphPad Software). Fractional saturation curves were produced with fractional saturation, Y, varying from 0–1.0. Y values are calculated by $(Y/(Y + (1 - Y)))$ and plotted against protein concentration. Data were converted to the Hill plot to analyse the degree of co-operativity in the binding events, characterized by the Hill coefficient (slope of the plot at $\log(\theta) = 0$). The Hill plot is constructed by plotting $\log\theta$ against log[protein], with θ defined as $(\theta = (Y/(1 - Y)))$. Dissociation coefficients (K_d) can also be extracted from the Hill plot as $K_d = 10^{X - \text{intercept}}$. Mean and standard error of the mean values are derived from at least three independent experiments.

Promoter activity assays

Promoter regions were cloned into the promoterless *lacZ* fusion plasmid, pRW50 [32]. Antitoxin genes were cloned into the pQE-80 derivative, pTA100 [5] for tight control of antitoxin expression. Construction is detailed above. Promoter activity assays were performed as described previously [45,46]. Briefly, *E. coli* DH5α were co-transformed with the *lacZ* reporter constructs pTRB483 (P_{rv1044}), pTRB484 ($P_{rv2827c}$) or pTRB486

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

Table 1 Crystallographic data collection and refinement statistics

	AbiEi Native	AbiEi Se-Peak	AbiEi Se-High Remote
PDB ID Code	6Y8Q	-	-
Number of crystals	3	2	1
Beamline	Diamond I03	Diamond I03	Diamond I03
Wavelength, Å	0.9763	0.9793	0.9641
Resolution range, Å	42.11–1.83 (1.86–1.83)[a]	42.58–2.14 (2.19–2.14)	53.57–2.17 (2.23–2.17)
Space group	P1	P1	P1
Unit cell, *a b c* (Å), $\alpha\ \beta\ \gamma$ (°)	34.24 80.85 122.17, 102.48 96.74 100.47	34.78 81.37 122.99, 101.72 97.18 101.16	34.85 81.38 123.00, 101.74 97.31 101.19
Total reflections	207 238(10 275)	443 813(13 873)	129 874(8557)
Unique reflections	106 620(5213)	69 714(4469)	65 917(4312)
Multiplicity	1.9	6.4	2.0
Completeness (%)	97.4 (96.1)	99.0 (97.1)	97.9 (91.9)
Mean I/sigma(I)	7.6	6.9	6.1
R_{merge}	0.038 (0.691)	0.169 (1.036)	0.080 (0.593)
R_{meas}	0.053 (0.977)	0.184 (1.260)	0.113 (0.839)
$CC_{1/2}$	0.999 (0.471)	0.991 (0.463)	0.992 (0.599)
R_{work}	0.1812 (0.2812)	-	-
R_{free}	0.2092 (0.3100)	-	-
No. of non-hydrogen			
atoms	7116	-	-
Macromolecules	6397	-	-
Ligands	62	-	-
Solvent	657	-	-
Protein Residues	769	-	-
RMSD (bonds, Å)	0.012	-	-
RMSD (angles, °)	1.32	-	-
Ramachandran			
favored (%)	98.68	-	-
Ramachandran			
allowed (%)	1.32	-	-
Ramachandran			
outliers (%)	0.00	-	-
Average B-factor	39.61	-	-
Macromolecules	39.04	-	-
Ligands	46.01	-	-
Solvent	44.60	-	-

(P_{abiEi}), and the IPTG-inducible pTA100-antitoxin plasmids pPF656 (Rv1044), pPF658 (Rv2827c) or pTRB481 (AbiEi). Transformants were re-seeded from overnight cultures and grown in 37°C at 200 rpm in LB supplemented with Tc, Sp, and with/without IPTG until mid-log phase, then 80 μl of cells were added to 120 μl master mix (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 36 mM β-mercaptoethanol, 166 μl/ml T7 lysozyme, 1.1 mg/ml ONPG, and 6.7% PopCulture Reagent (Merck Millipore)) in corresponding wells of a 96-well plate. This was then placed in a SPECTROstar Nano absorbance plate reader (BMG LABTECH) set to 30°C with shaking at 500 rpm, wherein OD_{600} and OD_{420} readings were taken every 90 s for

1 hour. Data analysis was performed in the MARS Data Analysis software package (BMG LABTECH). The kinetic OD_{420} readings were converted into the slope of OD_{420} over time (OD_{420}/min). These values were multiplied by 5000 and divided by the OD_{600} reading from the first time point to generate Miller Units (mU). Plotted data are the normalized mean and standard deviation obtained from three independent experiments.

Results

AbiEi-family antitoxins contain conserved structural features

We had previously hypothesized that there was structural similarity between the biochemically characterized antitoxin AbiEi from *S. agalactiae* [29] and the structurally characterized homolog, Rv2827c [30]. We sought to confirm structural and biochemical similarity between these two proteins, and within the broader COG5340 antitoxins. To begin, we solved a 1.83 Å structure of AbiEi by X-ray crystallography (Figure 1B,C and Table 1). There were four copies of AbiEi in the asymmetric unit, forming minor crystal contacts that are not predicted to be biologically relevant, and each copy contains minor variations in domain orientation, indicating some flexibility. Together with previous size exclusion chromatography data [29], we concluded that AbiEi is a 23 kDa monomer in solution.

AbiEi contains an N-terminal winged helix-turn-helix DNA-binding domain and a C-terminal antitoxin domain, connected by a short linker (Figure 1B). Mutagenesis studies have demonstrated that full-length AbiEi is required for negative autoregulation of the P_{abiE} promoter, whilst the C-terminal domain alone is sufficient for antitoxicity against the effects of AbiEii [15]. The N-terminal domain contains three α-helices, followed by three beta-strands forming an antiparallel sheet (Figure 1B). The C-terminal domain begins with a single α-helix that is separated from a six-helix bundle by a row of four β-strands, which themselves pair into parallel and antiparallel β-sheets (Figure 1B). One face of AbiEi is positively charged, and the reverse face is negatively charged (Figure 1C). The positive side corresponds with the site of positively charged sidechains distributed throughout the N-terminal and C-terminal domains, which have previously been shown to be vital for DNA-binding and autoregulation through mutagenesis studies [29]. When AbiEi is compared with Rv2827c, both are monomers and it is clear that the two antitoxins share the two-domain structure and charge features (Figure 1B–E).

When AbiEi and Rv2827c are aligned via the N-terminal winged helix-turn-helix domain, the respective C-terminal domains differ in position relative to the N-terminal domains by ∼65° (Figure 2A). We propose that these different poses captured in the crystal structures might reflect variable positions of the C-terminal domains potentially allowed by a linker joining the two domains. The stability of the B-factors for the subdomains AbiEi and Rv2827c, alongside lack of significant variation in the domain orientations within the asymmetric unit indicates a preferred state has been captured in the crystal. This however would require further analysis in solution. The extensive nature of the AbiEi charged surface, the requirement for the full AbiEi protein for autoregulation [15], and the presence of a flexible linker altogether indicate the full protein is needed for DNA interactions and DNA bending as per our previously proposed model [29].

When the N-terminal domain helices and C-terminal domains from AbiEi and Rv2827c are separated and structurally superposed, it is possible to see an approximate overlay between corresponding regions, with RMSDs of 3.04 Å for the N-terminal helices and 3.41 Å for the C-terminal domains (Figure 2B,C). The N-terminal domains have conserved positioning of key helices H2 and H3, which are used within winged helix-turn-helix domains for stabilization and DNA recognition, respectively [47] (Figure 2B). The C-terminal domain of AbiEi is the smaller of the two; performing a structure-based sequence alignment of AbiEi and Rv2827c shows that Rv2827c has an extended C-terminal domain 55 amino acids longer than AbiEi (Supplementary Figure S1). Despite this extension, AbiEi and Rv2827c share a conserved common core fold of unknown function (Figure 2C). When AbiEi was compared against the PDB to look for similar structures, using the DALI server [48], Rv2827c was the top hit followed by bacterial antibiotic-modifying adenylyltransferases (PDB codes 5KQJ, 4FO1), and a putative fungal NTase (PDB: 5UVD). These putative biochemical activities for AbiEi match well with the NTase activity of the cognate toxin AbiEii [15]. Overall, despite differing captured poses and discrepancy in size, AbiEi and Rv2827c are markedly similar in domain structure, fold and surface charge and are therefore structural homologs.

It has been shown that the AbiEi C-terminal domain is required for negative autoregulation and likely contributes to positive co-operativity through C-terminal domain interactions [29]. The cons-PPISP server [49] was used to highlight the residues most likely to be critical for protein–protein interactions for both AbiEi and

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

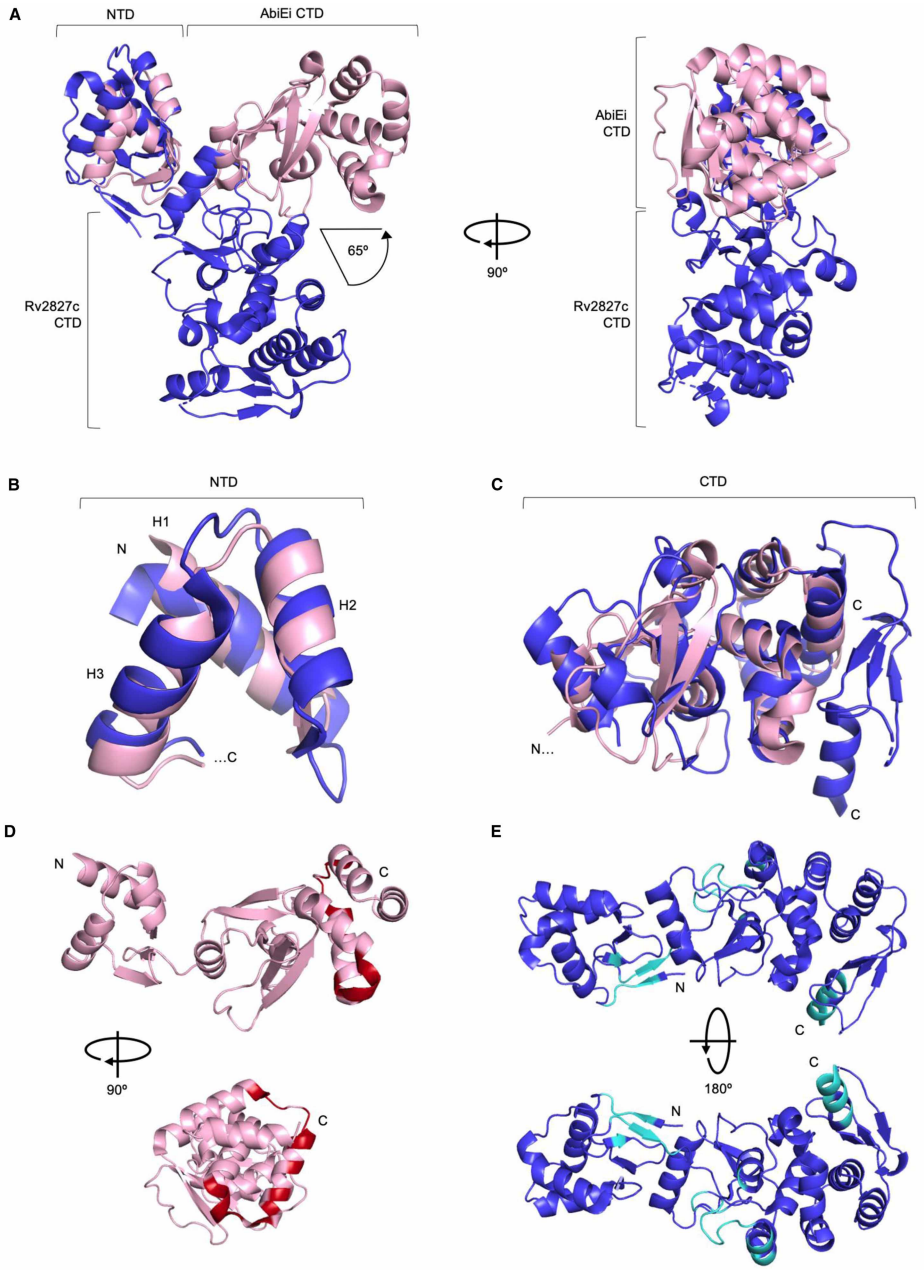


Figure 2. AbiEi and Rv2827c are structurally similar, but have been captured in different positions with differing predicted protein interaction interfaces.

(**A**) AbiEi (pink) and Rv2827c (blue) in cartoon representation, aligned via the N-terminal winged helix-turn-helix domains, shown as two orthogonal views. The positions of the C-terminal domains diverge at a 65° angle. (**B**) Close-up structural superposition of the isolated N-terminal helices of AbiEi and Rv2827c, colored as per (**A**). The three helices (H1–3) of the N-terminal winged helix-turn-helix domains align well. (**C**) Close-up structural superposition of the isolated C-terminal domains of AbiEi and Rv2827c, colored as per (**A**). The core secondary structural features of the C-terminal domains approximate to the same positions, but the Rv2827c C-terminal domain has additional features at the C-terminus. (**D**) AbiEi has C-terminal residues predicted to be involved in making protein–protein interactions, which might allow positive co-operativity in AbiEi monomer binding. AbiEi is in pink cartoon representation with identified interacting residues in red, and is shown in orthogonal views. (**E**) Rv2827c does not have an equivalent patch of C-terminal interacting residues. Rv2827c is in blue cartoon representation, with identified interacting residues in cyan, and is shown in 180° rotation. Residues were identified using the cons-PPISP server. Rv2827c PDB code: 1ZEL.

Rv2827c (Figure 2D,E). In the AbiEi monomer, 16 identified residues were clustered at the C-terminus, forming a putative site for interaction (Figure 2D). For Rv2827c, however, the diffuse scattering of 34 identified residues along the structure (Figure 2E) predicts that there may be no obvious interface for protein–protein interactions. This is reinforced by the different positioning of the CTD seen in Rv2827c (Figure 2A). These findings suggest that the interactions of AbiEi C-terminal domains could contribute to positive co-operativity in promoter binding, supporting our previously proposed model, whereas for Rv2827c, such interactions are unlikely to occur and indicate a different mechanism of DNA-binding and autoregulation.

Rv2827c binds two sets of inverted repeats

AbiEi binds to two 23 bp inverted repeats (IR1 and IR2) within the promoter of P_{abiEi}, which are separated by 3 bp [29] (Figure 3A). Examination of the upstream region of $P_{rv2827c}$ revealed two pairs of 23 bp inverted repeats within the region −1 to −131 bp from the *rv2827c* start codon, that also overlap the promoter (Figure 3A). These four repeats (IR1 to IR4) are arranged in tandem with a 4 bp gap between the two pairs of inverted repeats and a 13 bp gap between the repeats within each pair (Figure 3A). As P_{abiEi} repeats are separated by 3 bp and the repeats within pairs from $P_{rv2827c}$ are separated by 13 bp, it is possible the additional 10 bp accommodates binding of the larger C-terminal domains of Rv2827c (Supplementary Figure S1). Using the bacterial promoter prediction software CNNPromoter_b [50], the IR3–IR4 repeats were predicted to straddle a binding site for the primary *M. tuberculosis* sigma factor SigA [51]. As Rv2827c binding would sterically hinder sigma factor binding, in turn, this would prevent transcription of the operon by RNA polymerase. When IR1–IR4 sequences from $P_{rv2827c}$ were aligned with IR1–IR2 sequences from P_{abiEi}, the sequence similarity indicated possible conservation of binding sequence (Figure 3B). We therefore hypothesized that despite sharing low protein sequence identity (17.7%), Rv2827c might bind these $P_{rv2827c}$ inverted repeats similarly to AbiEi binding its cognate P_{abiEi} repeats.

The four $P_{rv2827c}$ 23 bp inverted repeats were first tested as two consecutive pairs, to allow a direct comparison to the arrangement of P_{abiEi} [29]. Analysis began with IR3–IR4, the pair overlapping the transcriptional start and therefore analogous to IR1–IR2 of P_{abiEi} (Figure 4A). Using electrophoretic mobility shift assays (EMSAs), Rv2827c was shown to bind both of the IR3–IR4 inverted repeats within the −1 to −71 region (Figure 4B). Sequential removal of the inverted repeats by mutating one, the other or both to polyC tracts reduced Rv2827c-DNA interaction to a single binding event (Figure 4C,D) or ablated binding completely

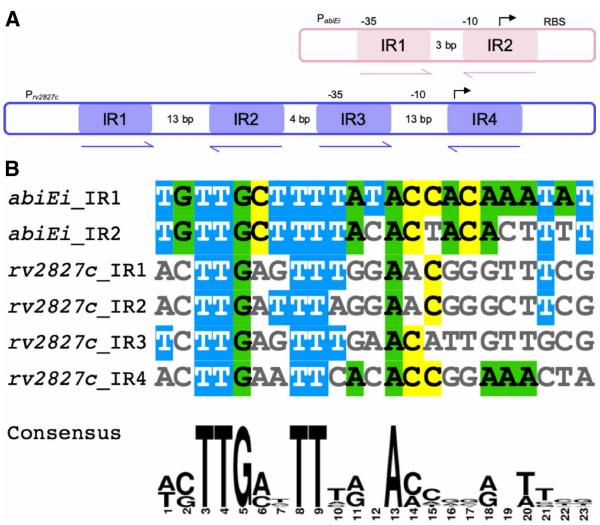


Figure 3. The *rv2827c–rv2826c* promoter has similar features but is more complex than the *abiE* promoter.
(**A**) Cartoon of the *abiE* and *rv2827c–rv2826c* promoters (pink and blue, respectively), showing the relative positions of the 23 bp inverted repeats (IRs). Putative transcriptional −35, −10 and start sites, along with ribosome binding sites (RBS), are indicated where possible. (**B**) Alignment of the six, 23 bp, IR sequences shows consensus sequences between the *abiE* and *rv2827c–rv2826c* promoters. The alignment was made using MView and the consensus was made using WebLogo.

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

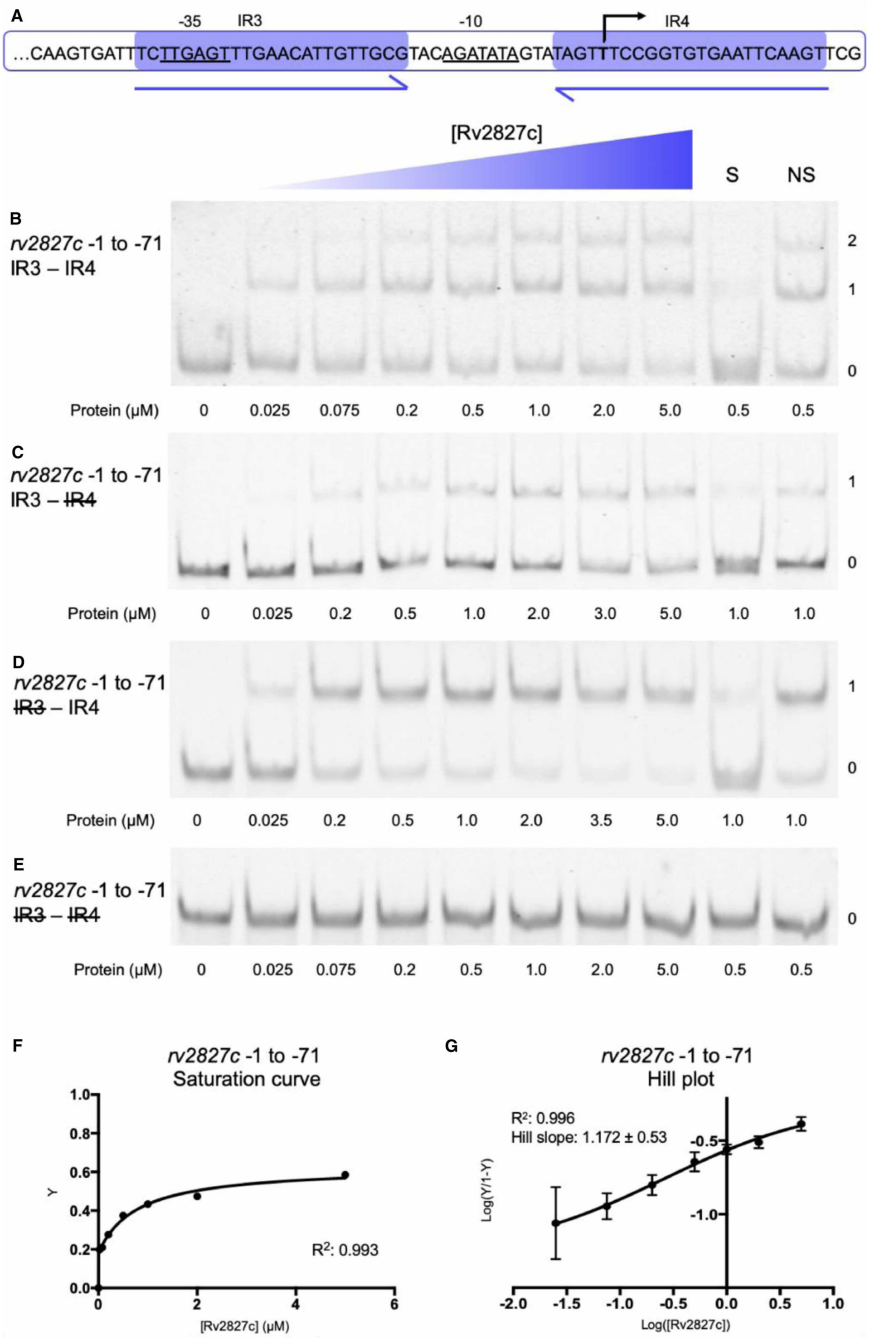


Figure 4. Rv2827c binds non-co-operatively to the IR3–IR4 region of the *rv2827c–rv2826c* promoter.

(**A**) Sequence level cartoon of the fluorescently labeled probe containing IR3–IR4, with −35, −10 and transcriptional start indicated. (**B**) Electrophoretic mobility shift assay (EMSA) of titrated Rv2827c with the probe in (**A**). (**C**) EMSA of titrated Rv2827c with the probe in (**A**) altered by replacing IR4 with polyC. (**D**) EMSA of titrated Rv2827c with the probe in (**A**) altered by replacing IR3 with polyC. (**E**) EMSA of titrated Rv2827c with the probe in (A) altered by replacing both IR3 and IR4 with polyC. For (**B**–**E**); protein concentrations are shown on each panel together with the binding events (0, 1 or 2); S — each experiment contained 100-fold excess of the specific unlabeled probe; NS — each experiment contained 100-fold excess of non-specific unlabeled probe; numbering −1 to −71 denotes the promoter region included in the probe, upstream of the translational start site in order to include all of IR4. (**F**) Fractional saturation curve plotted using the EMSA data of (**B**). (**G**) Hill plot using the EMSA data from (**B**). For (**F**) and (**G**), points are plotted from triplicate data and display mean values with standard error of the mean.

Biochemical Journal (2020) 477 2401–2419
https://doi.org/10.1042/BCJ20200368

(Figure 4E). Analysis of IR3–IR4 binding (Figure 4B,F) showed weak saturation of binding. The calculated Hill coefficient indicates that IR3–IR4 binding by Rv2827c is not co-operative (Figure 4G).

Similar results were obtained when testing the IR1–IR2 repeats within the −61 to −131 region of $P_{rv2827c}$ (Figure 5A–G). In this case, there was greater saturation of binding to IR1–IR2 (Figure 5F) than to IR3–IR4 (Figure 4F). The Hill coefficient surprisingly indicated weakly negative co-operativity in Rv2827c binding to IR1 and IR2 (Figure 5G), in comparison with the non-co-operative binding observed with IR3 and IR4 (Figure 4G). To allow direct comparison between model systems, we also performed the same assays with purified AbiEi and probes for P_{abiEi} (Supplementary Figure S2). This corroborated previous data [29] and under our experimental conditions, AbiEi bound more tightly to its cognate inverted repeats (Supplementary Figure S2F), than either Rv2827c binding to IR3–IR4 (Figure 4F) or IR1–IR2 (Figure 5F), and also demonstrated clear positive co-operativity (Supplementary Figure S2G).

Due to similarity of structure, functionality and cognate DNA inverted repeat sequences, we hypothesized that AbiEi and Rv2827c might bind their respective non-cognate promoter regions. However, AbiEi did not bind either pair of inverted repeats from $P_{rv2827c}$ (Supplementary Figure S3A,B). Similarly, Rv2827c did not bind IR1–IR2 of P_{abiEi} (Supplementary Figure S3C).

Rv2827c binds with negative co-operativity

Having investigated the two sets of $P_{rv2827c}$ inverted repeats independently, a full-length probe covering the $P_{rv2827c}$ region −1 to −131 was generated to examine the interaction of Rv2827c protein with all four inverted repeats. Using EMSAs, four distinct protein-bound DNA species were observed, indicating that all four inverted repeats can be bound simultaneously by Rv2827c (Figure 6A). The four binding sites did not fully saturate (Figure 6B), and the Hill coefficient confirmed negatively co-operative binding of Rv2827c across these four inverted repeats (Figure 6C). Displaying the saturation curve data on a semi-log scale highlights breaks and multiple distinct gradients in the binding curve, eluding to multiple individual binding events (Figure 6D). Negatively co-operative binding by Rv2827c to $P_{rv2827c}$ contrasts with the positive co-operativity observed for AbiEi binding to P_{abiEi} [29].

Our earlier data using mutant probes provided insight into how Rv2827c binds to individual repeats (Figures 4C,D and 5C,D). These were used to calculate the binding affinity of Rv2827c for each individual IR sequence, with IR1 most tightly bound (K_d of 0.0205 μM), closely followed by IR4 (K_d of 0.121 μM), then IR2 (K_d of 0.862 μM), and finally IR3 (K_d of 11.0 μM) (Figure 6E–M). This descending affinity series creates a wide range of concentrations across which negative autoregulation by Rv2827c can occur. These data demonstrate the same core principles of promoter binding are used by both AbiEi and Rv2827c, but that these have been employed evolutionarily for differing modes of regulation.

Rv1044 is a DNA-binding protein, but fails to recognize the cognate promoter

Whilst it had not been possible to identify inverted repeats within the *rv1044–rv1045* locus, two, slightly overlapping, 70 bp probes were generated to cover the 131 bp region upstream of the *rv1044* translational start site, and used to test Rv1044 binding (Supplementary Figure S4A,B). No binding event was observed with either probe (Supplementary Figure S4A,B). Nevertheless, we wanted to test whether Rv1044 was competent for DNA-binding and so cross-reacted Rv1044 with the two probes covering IR3–IR4 and IR1–IR2 of $P_{rv2827c}$, and the probe containing IR1–IR2 of P_{abiEi} (Supplementary Figure S4C–E). No binding was observed for either of the $P_{rv2827c}$ probes (Supplementary Figure S4C,D), but curiously, Rv1044 bound the inverted repeats of P_{abiEi} (Supplementary Figure S4E). Rv1044 bound more weakly than AbiEi to P_{abiEi} IR1–IR2 (Supplementary Figure S4F), and showed no co-operativity (Supplementary Figure S4G). This demonstrates that Rv1044 can bind DNA in a sequence-specific manner, and so we looked for potential targets in the *M. tuberculosis* H37Rv genome. The *abiE* IR sequences align with numerous positions in the *M. tuberculosis* genome but not upstream of any of the DUF1814 TA systems. This may indicate a potential role for Rv1044 in regulating genes outside of the *rv1044–rv1045* operon, as has been shown in other TA systems whereby antitoxins influence gene expression in biofilm formation pathways [52–54]. A further study will be needed to fully explore any potential regulatory role of Rv1044.

Rv2827c negatively autoregulates Rv2827c–Rv2826c expression

Having shown a structural similarity between the two antitoxins, we next sought to test whether the COG5340 proteins from *M. tuberculosis* could function as autoregulators, like characterized AbiEi [29]. AbiEi negatively

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

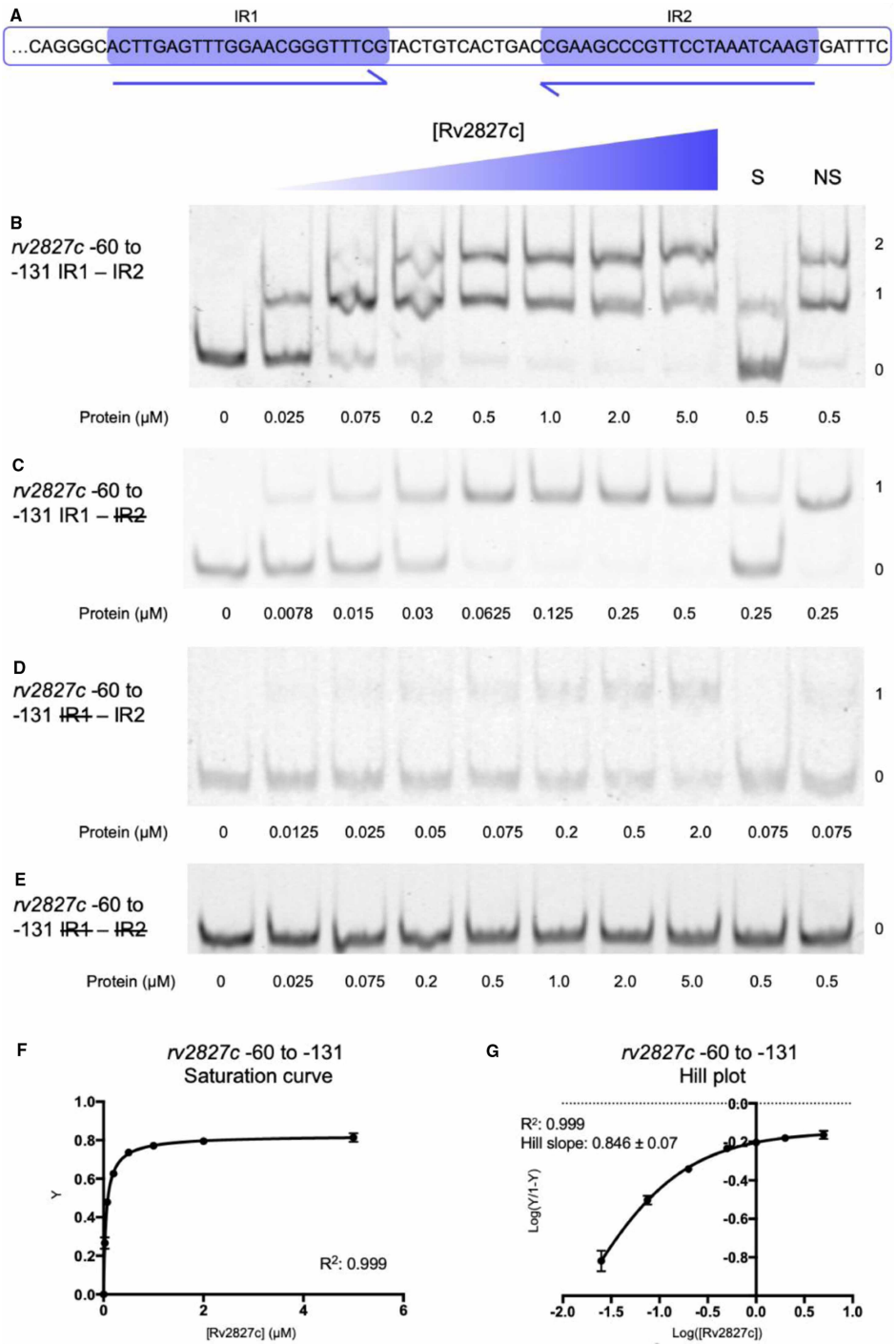


Figure 5. Rv2827c binds with weak negative co-operativity to the IR1–IR2 region of the *rv2827c–rv2826c* promoter.
(**A**) Sequence level cartoon of the fluorescently labeled probe containing IR1–IR2. (**B**) Electrophoretic mobility shift assay (EMSA) of titrated Rv2827c with the probe in (**A**). (**C**) EMSA of titrated Rv2827c with the probe in (**A**) altered by replacing IR2 with polyC. (**D**) EMSA of titrated Rv2827c with the probe in (**A**) altered by replacing IR1 with polyC. (**E**) EMSA of titrated Rv2827c with the probe in (**A**) altered by replacing both IR1 and IR2 with polyC. For (**B**–**E**); protein concentrations are shown on each panel together with the binding events (0, 1 or 2); S — each experiment contained 100-fold excess of the specific unlabeled probe; NS — each experiment contained 100-fold excess of non-specific unlabeled probe; numbering −60 to −131 denotes the promoter region included in the probe. (**F**) Fractional saturation curve plotted using the EMSA data of (**B**). (**G**) Hill plot using the EMSA data from (**B**). For (**F**) and (**G**), points are plotted from triplicate data and display mean values with standard error of the mean.

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

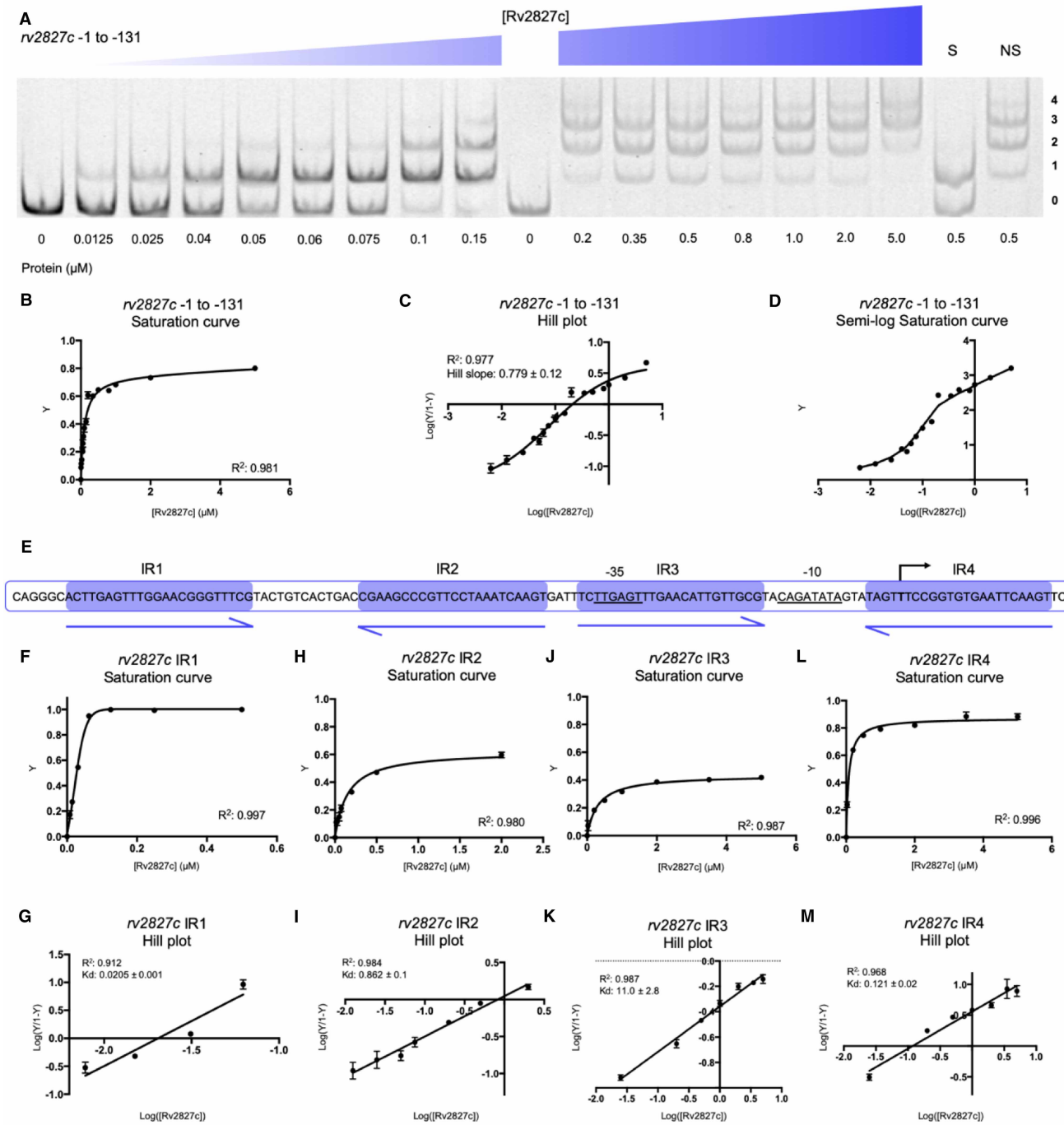


Figure 6. Rv2827c binds with negative co-operativity to the full *rv2827c–rv2826c* promoter.

(**A**) EMSA of titrated Rv2827c with a probe covering from −1 to −131 of the *rv2827c–rv2826c* promoter, covering IR1 to IR4. The titration was performed across two EMSA gels, with an additional zero protein lane included in the second gel for normalization. Protein concentrations are shown below each gel together with the binding events (0, 1, 2, 3 or 4); S — each experiment contained 100-fold excess of the specific unlabeled probe; NS — each experiment contained 100-fold excess of non-specific unlabeled probe. (**B**) Fractional saturation curve plotted using the EMSA data of (**A**). (**C**) Hill plot using the EMSA data from (**A**). (**D**) Semi-log saturation curve plotted using the EMSA data of (**A**), showing distinct breaks in the binding curve, in accordance with the multiple binding sites contained in the probe. (**E**) Sequence level cartoon of the fluorescently labeled probe containing *rv2827c–rv2826c* −1 to −131. (**F**–**M**) Saturation curves (**F**,**H**,**J**,**L**) and Hill plots (**G**,**I**,**K**,**M**) for each IR calculated using individual IR data gathered using mutant probes (Figures 4C, D and 5C,D). For (**B**–**D**) and (**F**–**M**), points are plotted from triplicate data and display mean values with standard error of the mean.

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

autoregulates expression from the P_{abiEi} promoter [29]. To examine whether Rv2827c and also the second *M. tuberculosis* COG5340 protein, Rv1044, also perform negative autoregulation, we first cloned the 500 bp region upstream of each respective translational start site into a promoterless *lacZ*-reporter plasmid [32]. For comparison, the equivalent P_{abiEi}–reporter, containing the previously identified promoter region identified in the upstream 99 bp [15,29] was also tested. Both P_{abiEi} and $P_{rv2827c}$ reporters yielded expression of LacZ, but P_{rv1044} did not (Figure 7A). The two active reporter constructs P_{abiEi} and $P_{rv2827c}$, were then paired with inducible plasmids for expression of the cognate antitoxins, and LacZ levels were determined with and without antitoxin induction (Figure 7B). When compared with the uninduced controls, both antitoxins negatively autoregulated expression from their cognate promoters (Figure 7B) demonstrating that Rv2827c and AbiEi share not only a common structure, but also a common negative autoregulatory function.

Discussion

In this study we present the crystal structure of *S. agalactiae* AbiEi, which was the first type IV TA system antitoxin shown to be capable of transcriptional autoregulation through promoter binding [29]. Further to this, we have demonstrated the autoregulatory capacity of the related Rv2827c antitoxin, a protein of known structure [30]. Whilst AbiEi is a structural homolog of the Rv2827c antitoxin, and both share similar promoter architectures, they have distinct differences in their size and captured domain orientations. We also show that negative autoregulation of the $P_{rv2827c}$ promoter operates via negatively co-operative interactions.

Despite the low shared sequence similarity seen for the COG5340 antitoxins investigated (AbiEi and Rv2827c — 17.7%; AbiEi and Rv1044 — 21.2%; Rv2827c and Rv1044 — 24.5%), we have demonstrated structural conservation across species. As sequences diverge, structure is conserved (Figures 1, 2), which maintains the shared functionality of these antitoxins, for instance, DNA-binding (Figures 4–7 and Supplementary Figures S2–S4). Interestingly, sequence variation of the NTD, alongside differing promoter architectures, has resulted in at least three variations of antitoxicity. AbiEi and Rv2827c both autoregulate their own operons, albeit with contrasting types of co-operativity. Rv1044, however, may regulate genes elsewhere in the *M. tuberculosis* genome, given the lack of affinity to the *rv1044* upstream region tested (Supplementary Figure S4A,B)

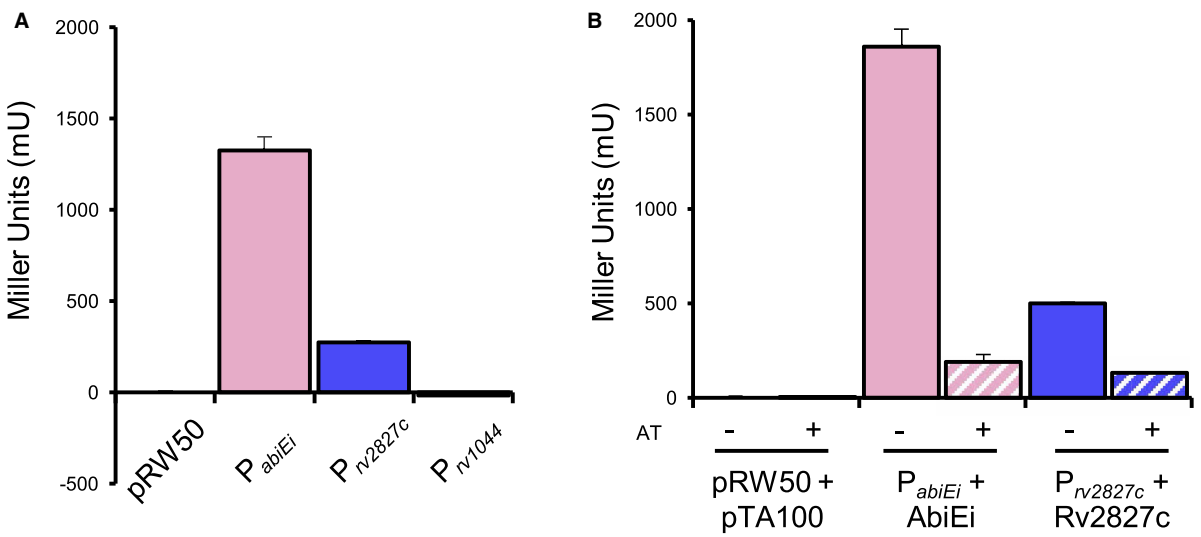


Figure 7. Rv2827c–Rv2826c is a negatively autoregulating system in *E. coli*.

(**A**) Promoter activity from upstream promoter regions of *abiE* (99 bp), and *rv2827c–rv2826c* and *rv1044–rv1045* (500 bp for both) detected using *lacZ* transcriptional fusions. Both the *abiE* and *rv2827c–rv2826c* constructs are active, but the *rv1044–rv1045* construct is not. Plotted data are normalized to the vector-only control. (**B**) Autoregulation of promoter activity by antitoxins. LacZ activity was measured from the *abiE* and *rv2827c–rv2826c* constructs with or without induction of the cognate antitoxin (AT, ±IPTG). Both AbiE and Rv2827c negatively autoregulate expression. Plotted data are normalized to the uninduced vector-only control. All data (**A–B**) are plotted as the means of triplicate data, and error bars show standard deviations from the mean.

and absence of identifiable inverted repeats, but apparent DNA-binding capabilities (Supplementary Figure S4E–G). Further analysis will be required to identify a functional promoter for the *rv1044–rv1045* operon and confirm any potential regulatory function of Rv1044. The antitoxic CTDs have a common core fold that are predicted to have NTase activity based on structure-based functional searches [30]. Therefore, the antitoxic mechanism is likely conserved, despite low sequence similarity within these domains (Supplementary Figure S1). As protein sequences will be tuned to the needs of the organism, we have shown a correspondingly differential pattern of residues for protein–protein interactions (Figure 2D,E) which, alongside the different CTD positions captured (Figure 2A), may contribute to individual autoregulation requirements. Our previous model predicted AbiEi C-terminal domain interactions promote positive co-operative binding and result in DNA-bending [29], however this does not appear to apply to Rv2827c. Our proposed model (Figure 8) implies

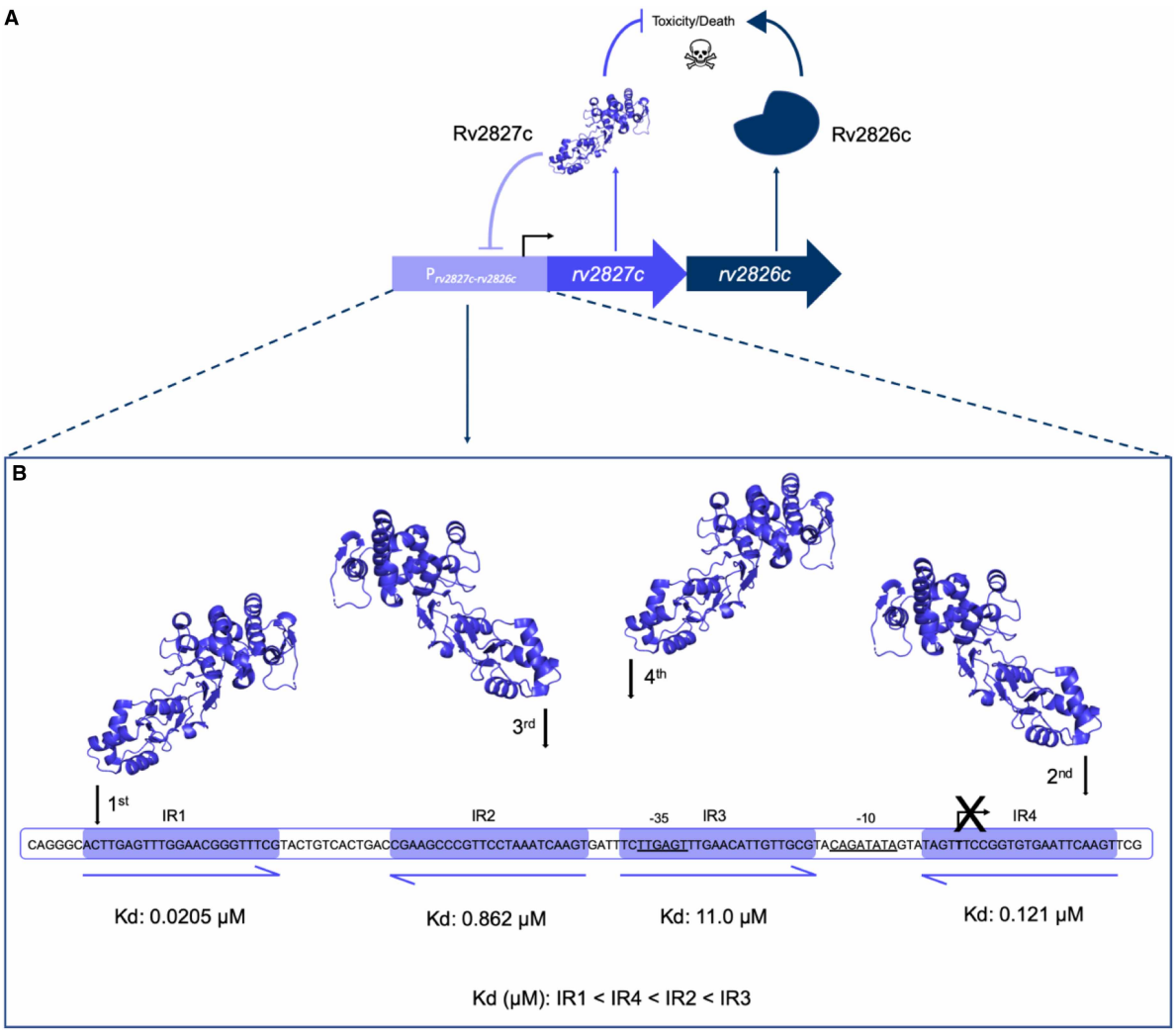


Figure 8. Proposed model for negative autoregulation caused by Rv2827c binding to the four *rv2827c–rv2826c* promoter inverted repeats.

(**A**) Schematic representation of the putative rv2827c–rv2826c type IV toxin-antitoxin system. Model shows both *rv2827c* and *rv2826c* being translated into the antagonistic antitoxin and toxin protein pair, respectively. The antitoxin, Rv2827c, has a second function and binds to the *rv2827c–rv2826c* promoter, negatively autoregulating the operon. (**B**) An order of binding is created by the distinct affinity values for the inverted repeats represented in the sequence level cartoon, calculated from individual IR data gathered using mutant probes (Figures 4C,D and 5C, D). Rv2827c binds negatively co-operatively, initially to IR1 (0.0205 μM) followed by IR4 (0.121 μM), IR2 (0.862 μM) and finally IR3 (11.0 μM).

Biochemical Journal (2020) 477 2401–2419
https://doi.org/10.1042/BCJ20200368

a possible lack of protein–protein interactions supported by predicted interaction interfaces (Figure 2D,E), while not ruling out the potential for steric restriction. Rather, promoter inverted repeat sequence 'tuning' (Figure 3) contributes to the negatively co-operative interaction via descending affinities.

Promoters of *M. tuberculosis* are known to be more complex than those of *E. coli*; they can stretch to 2000 bp from the start site and lack canonical elements such as the conserved −35 sequence [55–57]. Transcriptional regulation is complicated further when considering the vast number of sigma factors [58] and environmentally responsive transcription factors [59] present in *M. tuberculosis*, allowing for greater promoter sequence variation. The −10 sequence for *rv2827c–rv2826c* is a predicted recognition site for principle *M. tuberculosis* sigma factor SigA, which is usually maintained at a constant level for cellular 'housekeeping' [51]. SigA also has a role in host-pathogen interactions, controlling growth rates during macrophage infection [60] and regulating virulence genes through both constitutive and up-regulated expression [61–63]. Deletions of *rv2827c* cause a growth defect [25,26], suggesting SigA drives expression and that there is potential for output to be tuned by SigA and Rv2827c levels according to environmental cues. Previous reports on the type IV antitoxin CbeA demonstrate a positive effect on cytoskeletal bundling alongside antitoxicity and the ability to counteract chemical inhibitors of cytoskeletal polymerisation [13]. One study has shown Rv2827c up-regulation in response to isoniazid and rifampicin treatment, albeit as part of more general TA system up-regulation [64].

The IR conservation between P_{abiEi} and $P_{rv2827c}$ (Figure 3) suggested that autoregulation may also occur in a biochemically comparable manner between the two. However, Rv2827c bound the pairs of inverted repeats with either no co-operativity (Figure 4), or weakly negative co-operativity (Figure 5). There was clear negative co-operativity when all four inverted repeats were tested (Figure 6). Analyzing each inverted repeat independently by mutational studies identified significant differences between the Rv2827c-IR dissociation constants (Figure 6E–M). These data have allowed us to propose a model for the regulation of *rv2827c–rv2826c* (Figure 8). As *rv2827c* is needed for normal growth, this suggests that *rv2826c* encodes a toxin capable of causing growth defects [25,65], which is antagonized by Rv2827c (Figure 8A). Expression of *rv2827c–rv2826c* is negatively autoregulated by Rv2827c, and this is made possible by sequential binding of Rv2827c to the four IR sequences, in order as determined by binding affinity (Figure 8B). Given the high concentration of Rv2827c required to saturate the lower affinity site IR3 (Figures 6J,K and 8B), mimicking the mutational analysis performed here in promoter activity studies would provide useful insight into the function of each IR sequence. These binding events have apparent negative co-operativity, likely due to the variations in IR sequences creating a series of binding steps with ever-decreasing affinity. To better understand these negatively co-operative interactions further experiments are required, exploring the role of the Rv2827c CTD and larger inverted repeat spacers, akin to previous work on AbiEi [29].

Negative co-operativity was an unexpected result given the structural similarities between the N-terminal domains of AbiEi and Rv2827c (Figure 2B), and the similarities of their respective promoter architectures (Figure 3). Examples of negative and positive co-operativity have been found in equal abundance across all organisms [66,67]. Positive co-operativity leads to rapid saturation at a defined, short range of concentrations as seen for *abiE* [29]. In contrast, negative co-operativity of Rv2827c binding would be expected to generate a relatively delayed response, working across a greater range of Rv2827c concentrations [67–69]. This variability in tuning according to concentration could in turn relate to the relative potency of the toxins and dosage required to have an effect in their cognate hosts. This variation is evident when comparing saturation curves of AbiEi and Rv2827c to their cognate full-length promoters (Figure 6B and Supplementary Figure S2F). Compared with positive co-operativity, there is relatively little information on the presence of negatively co-operative TA-promoter interactions. However, clear evidence supports weaker binding of un-complexed type II antitoxins [52,70] when compared with the conditionally co-operative binding of TA complexes [28,52,71,72]. It is noteworthy that unlike many type II antitoxins, AbiEi and Rv2827c are fully folded and stable, and also no conditionally co-operative response was seen for AbiE, and so the conditional model proposed for many type II systems likely does not apply [15].

This study has shown that the similar structures and promoter architectures between AbiEi, Rv2827c (and indeed Rv1044) have been co-opted to form different regulatory modules. A greater understanding of how these nuances of regulation are applied in the cognate hosts may provide greater insight into the control of bacterial growth. Understanding these systems and how they regulate bacterial behavior is thereby an important step in developing a means to control TA systems towards utilizing them for their potential therapeutic value.

Data Availability

The crystal structure of AbiEi has been deposited in the Protein Data Bank under accession number 6Y8Q.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by a Springboard Award from the Academy of Medical Sciences (SBF002\1104) [I.N. B., B.U., T.R.B.], a BBSRC NLD Doctoral Training Partnership studentship (BB/M011186/1) [I.N.B.], a University of Otago Doctoral Scholarship [H.G.H], a University of Otago Research Grant [P.C.F], School of Biomedical Sciences Bequest Funds, University of Otago [P.C.F] and the Matariki Network of Universities (MNU) [P.C.F].

Author Contributions

Conceptualization, all authors; Investigation, I.N.B., B.U., H.H.; Writing, all authors; Funding acquisition, P.C.F., T.R.B.; Supervision, P.C.F., T.R.B.

Acknowledgements

We thank Koen Semeijn and Ron Dy for initial plasmid construction and preliminary testing.

Abbreviations

CTDs, C-terminal domains; EMSAs, electrophoretic mobility shift assays; IPTG, isopropyl-β-D-thiogalactopyranoside; LIC, Ligation Independent Cloning; NTase, nucleotidyltransferase.

References

1 Norton, J.P. and Mulvey, M.A. (2012) Toxin-antitoxin systems are important for niche-specific colonization and stress resistance of uropathogenic *Escherichia coli*. *PLoS Pathog.* **8**, e1002954 https://doi.org/10.1371/journal.ppat.1002954
2 Van Acker, H. and Coenye, T. (2017) The role of reactive oxygen species in antibiotic-mediated killing of bacteria. *Trends Microbiol.* **25**, 456–466 https://doi.org/10.1016/j.tim.2016.12.008
3 Helaine, S., Cheverton, A.M., Watson, K.G., Faure, L.M., Matthews, S.A. and Holden, D.W. (2014) Internalization of salmonella by macrophages induces formation of nonreplicating persisters. *Science* **343**, 204–208 https://doi.org/10.1126/science.1244705
4 Page, R. and Peti, W. (2016) Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.* **12**, 208–214 https://doi.org/10.1038/nchembio.2044
5 Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S. and Salmond, G.P. (2009) The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl Acad. Sci. U.S.A.* **106**, 894–899 https://doi.org/10.1073/pnas.0808832106
6 Fraikin, N., Goormaghtigh, F. and Van Melderen, L. (2020) Type II toxin-antitoxin systems: evolution and revolutions. *J. Bacteriol.* **202**, e00763-19 https://doi.org/10.1128/JB.00763-19
7 Szekeres, S., Dauti, M., Wilde, C., Mazel, D. and Rowe-Magnus, D.A. (2007) Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol. Microbiol.* **63**, 1588–1605 https://doi.org/10.1111/j.1365-2958.2007.05613.x
8 Wozniak, R.A.F. and Waldor, M.K. (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* **8**, 552–563 https://doi.org/10.1038/nrmicro2382
9 Deter, H., Jensen, R., Mather, W. and Butzin, N. (2017) Mechanisms for differential protein production in toxin–antitoxin systems. *Toxins* **9**, 211 https://doi.org/10.3390/toxins9070211
10 Hall, A.M., Gollan, B. and Helaine, S. (2017) Toxin–antitoxin systems: reversible toxicity. *Curr. Opin. Microbiol.* **36**, 102–110 https://doi.org/10.1016/j.mib.2017.02.003
11 Labrie, S.J., Samson, J.E. and Moineau, S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 https://doi.org/10.1038/nrmicro2315
12 Chopin, M.C., Chopin, A. and Bidnenko, E. (2005) Phage abortive infection in lactococci: variations on a theme. *Curr. Opin. Microbiol.* **8**, 473–479 https://doi.org/10.1016/j.mib.2005.06.006
13 Masuda, H., Tan, Q., Awano, N., Wu, K.-P. and Inouye, M. (2012) Yeeu enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. *Mol. Microbiol.* **84**, 979–989 https://doi.org/10.1111/j.1365-2958.2012.08068.x
14 Masuda, H., Tan, Q., Awano, N., Yamaguchi, Y. and Inouye, M. (2012) A novel membrane-bound toxin for cell division, CptA (YgfX), inhibits polymerization of cytoskeleton proteins, FtsZ and MreB, in *Escherichia coli*. *FEMS Microbiol. Lett.* **328**, 174–181 https://doi.org/10.1111/j.1574-6968.2012.02496.x
15 Dy, R.L., Przybilski, R., Semeijn, K., Salmond, G.P.C. and Fineran, P.C. (2014) A widespread bacteriophage abortive infection system functions through a type IV toxin-antitoxin mechanism. *Nucleic Acids Res.* **42**, 4590–4605 https://doi.org/10.1093/nar/gkt1419
16 Blower, T.R., Pei, X.Y., Short, F.L., Fineran, P.C., Humphreys, D.P., Luisi, B.F. et al. (2011) A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat. Struct. Mol. Biol.* **18**, 185–190 https://doi.org/10.1038/nsmb.1981
17 Emond, E., Dion, E., Walker, S.A., Vedamuthu, E.R., Kondo, J.K. and Moineau, S. (1998) AbiQ, an abortive infection mechanism from *Lactococcus lactis*. *Appl. Environ. Microbiol.* **64**, 4748–4756 https://doi.org/10.1128/AEM.64.12.4748-4756.1998

Biochemical Journal (2020) **477** 2401–2419
https://doi.org/10.1042/BCJ20200368

18 Samson, J.E., Spinelli, S., Cambillau, C. and Moineau, S. (2013) Structure and activity of AbiQ, a lactococcal endoribonuclease belonging to the type III toxin-antitoxin system. *Mol. Microbiol.* **87**, 756–768 https://doi.org/10.1111/mmi.12129

19 Pecota, D.C. and Wood, T.K. (1996) Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *J. Bacteriol.* **178**, 2044–2050 https://doi.org/10.1128/JB.178.7.2044-2050.1996

20 Hazan, R. and Engelberg-Kulka, H. (2004) *Escherichia coli mazEF*-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Mol. Genet. Genomics* **272**, 227–234 https://doi.org/10.1007/s00438-004-1048-y

21 Keren, I., Minami, S., Rubin, E. and Lewis, K. (2011) Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *MBio* **2**, e00100-11 https://doi.org/10.1128/mBio.00100-11

22 Sala, A., Bordes, P. and Genevaux, P. (2014) Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins* **6**, 1002–1020 https://doi.org/10.3390/toxins6031002

23 Slayden, R.A., Dawson, C.C. and Cummings, J.E. (2018) Toxin–antitoxin systems and regulatory mechanisms in *Mycobacterium tuberculosis*. *Pathog. Dis.* **76** https://doi.org/10.1093/femspd/fty039

24 Akarsu, H., Bordes, P., Mansour, M., Bigot, D.J., Genevaux, P. and Falquet, L. (2019) TASmania: a bacterial toxin-antitoxin systems database. *PLoS Comput. Biol.* **15**, e1006946 https://doi.org/10.1371/journal.pcbi.1006946

25 Sassetti, C.M., Boyd, D.H. and Rubin, E.J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 https://doi.org/10.1046/j.1365-2958.2003.03425.x

26 Dejesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C. et al. (2017) Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *MBio* **8**, e02133-16 https://doi.org/10.1128/mBio.02133-16

27 Gerdes, K., Christensen, S.K. and Lobner-Olesen, A. (2005) Prokaryotic toxin-antitoxin stress response loci. *Nat. Rev. Microbiol.* **3**, 371–382 https://doi.org/10.1038/nrmicro1147

28 Overgaard, M., Borch, J., Jørgensen, M.G. and Gerdes, K. (2008) Messenger RNA interferase RelE controls *relBE* transcription by conditional cooperativity. *Mol. Microbiol.* **69**, 841–857 https://doi.org/10.1111/j.1365-2958.2008.06313.x

29 Hampton, H.G., Jackson, S.A., Fagerlund, R.D., Vogel, A.I.M., Dy, R.L., Blower, T.R. et al. (2018) AbiEi binds cooperatively to the type IV *abiE* toxin–antitoxin operator via a positively-charged surface and causes DNA bending and negative autoregulation. *J. Mol. Biol.* **430**, 1141–1156 https://doi.org/10.1016/j.jmb.2018.02.022

30 Janowski, R., Panjikar, S., Eddine, A.N., Kaufmann, S.H.E. and Weiss, M.S. (2009) Structural analysis reveals DNA binding properties of Rv2827c, a hypothetical protein from *Mycobacterium tuberculosis*. *J. Struct. Funct. Genomics* **10**, 137–150 https://doi.org/10.1007/s10969-009-9060-4

31 Aslanidis, C. and de Jong, P.J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* **18**, 6069–6074 https://doi.org/10.1093/nar/18.20.6069

32 Lodge, J., Fear, J., Busby, S., Gunasekaran, P. and Kamini, N.R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett.* **95**, 271–276 https://doi.org/10.1111/j.1574-6968.1992.tb05378.x

33 Kabsch, W. (2010) XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 https://doi.org/10.1107/S0907444909047337

34 Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 133–144 https://doi.org/10.1107/S0907444909047374

35 Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R. et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 https://doi.org/10.1107/S0907444910045749

36 Sheldrick, G.M. (2008) IUCr, A short history of *SHELX*. *Acta Crystallogr. Sect. A Found. Crystallogr.* **64**, 112–122 https://doi.org/10.1107/S0108767307043930

37 Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F. et al. (2004) *REFMAC* 5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2184–2195 https://doi.org/10.1107/S0907444904023510

38 Cowtan, K. (2006) IUCr, The buccaneer software for automated model building. 1. tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 1002–1011 https://doi.org/10.1107/S0907444906022116

39 Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 https://doi.org/10.1107/S0907444909052925

40 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* **60**, 2126–2132 https://doi.org/10.1107/S0907444904019158

41 Gore, S., Velankar, S. and Kleywegt, G.J. (2012) Implementing an X-ray validation pipeline for the protein data bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 478–483 https://doi.org/10.1107/S0907444911050359

42 Pei, J., Kim, B.-H. and Grishin, N.V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300 https://doi.org/10.1093/nar/gkn072

43 Madeira, F., mi Park, Y., Lee, J., Buso, N., Gur, T., Madhusoodanan, N. et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 https://doi.org/10.1093/nar/gkz268

44 Crooks, G.E., Hon, G., Chandonia, J.-M. and Brenner, S.E. (2004) Weblogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 https://doi.org/10.1101/gr.849004

45 Schaefer, J., Jovanovic, G., Kotta-Loizou, I. and Buck, M. (2016) A data comparison between a traditional and the single-step β-galactosidase assay. *Data Br.* **8**, 350–352 https://doi.org/10.1016/j.dib.2016.05.063

46 Schaefer, J., Jovanovic, G., Kotta-Loizou, I. and Buck, M. (2016) Single-step method for β-galactosidase assays in *Escherichia coli* using a 96-well microplate reader. *Anal. Biochem.* **503**, 56–57 https://doi.org/10.1016/j.ab.2016.03.017

47 Gajiwala, K.S. and Burley, S.K. (2000) Winged helix proteins. *Curr. Opin. Struct. Biol.* **10**, 110–116 https://doi.org/10.1016/S0959-440X(99)00057-3

48 Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 https://doi.org/10.1006/jmbi.1993.1489

49 Chen, H. and Zhou, H.-X. (2005) Prediction of interface residues in protein-protein complexes by a consensus neural network method: Test against NMR data. *Proteins Struct. Funct. Bioinformatics* **61**, 21–35 https://doi.org/10.1002/prot.20514

50 Umarov, R.K. and Solovyev, V.V. (2017) Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks. *PLoS ONE* **12**, e0171410 https://doi.org/10.1371/journal.pone.0171410
51 Rodrigue, S., Brodeur, J., Jacques, P.-E., Gervais, A.L., Brzezinski, R. and Gaudreau, L. (2007) Identification of mycobacterial sigma factor binding sites by chromatin immunoprecipitation assays. *J. Bacteriol.* **189**, 1505–1513 https://doi.org/10.1128/JB.01371-06
52 Merfa, M.V., Niza, B., Takita, M.A. and De Souza, A.A. (2016) The MqsRA toxin-antitoxin system from *Xylella fastidiosa* plays a key role in bacterial fitness, pathogenicity, and persister cell formation. *Front. Microbiol.* **7**, 904 https://doi.org/10.3389/fmicb.2016.00904
53 Soo, V.W.C. and Wood, T.K. (2013) Antitoxin MqsA represses curli formation through the master biofilm regulator CsgD. *Sci. Rep.* **3**, 3186 https://doi.org/10.1038/srep03186
54 Wang, X. and Wood, T.K. (2011) Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl. Environ. Microbiol.* **77**, 5577–5583 https://doi.org/10.1128/AEM.05068-11
55 Agarwal, N. and Tyagi, A.K. (2006) Mycobacterial transcriptional signals: requirements for recognition by RNA polymerase and optimal transcriptional activity. *Nucleic Acids Res.* **34**, 4245–4257 https://doi.org/10.1093/nar/gkl521
56 Shell, S.S., Wang, J., Lapierre, P., Mir, M., Chase, M.R., Pyle, M.M. et al. (2015) Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. *PLoS Genet.* **11**, e1005641 https://doi.org/10.1371/journal.pgen.1005641
57 Cortes, T., Schubert, O.T., Rose, G., Arnvig, K.B., Comas, I., Aebersold, R. et al. (2013) Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis*. *Cell Rep.* **5**, 1121–1131 https://doi.org/10.1016/j.celrep.2013.10.031
58 Sachdeva, P., Misra, R., Tyagi, A.K. and Singh, Y. (2010) The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J.* **277**, 605–626 https://doi.org/10.1111/j.1742-4658.2009.07479.x
59 Rustad, T.R., Minch, K.J., Ma, S., Winkler, J.K., Hobbs, S., Hickey, M. et al. (2014) Mapping and manipulating the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biol.* **15**, 502 https://doi.org/10.1186/s13059-014-0502-3
60 Wu, S., Howard, S.T., Lakey, D.L., Kipnis, A., Samten, B., Safi, H. et al. (2004) The principal sigma factor sigA mediates enhanced growth of *Mycobacterium tuberculosis in vivo*. *Mol. Microbiol.* **51**, 1551–1562 https://doi.org/10.1111/j.1365-2958.2003.03922.x
61 Vallecillo, A.J. and Espitia, C. (2009) Expression of *Mycobacterium tuberculosis pe_pgrs33* is repressed during stationary phase and stress conditions, and its transcription is mediated by sigma factor A. *Microb. Pathog.* **46**, 119–127 https://doi.org/10.1016/j.micpath.2008.11.003
62 Bagchi, G., Chauhan, S., Sharma, D. and Tyagi, J.S. (2005) Transcription and autoregulation of the *rv3134c-devR-devS* operon of *Mycobacterium tuberculosis*. *Microbiology.* **151**, 4045–4053 https://doi.org/10.1099/mic.0.28333-0
63 Rengarajan, J., Bloom, B.R. and Rubin, E.J. (2005) Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proc. Natl Acad. Sci. U.S.A.* **102**, 8327–8332 https://doi.org/10.1073/pnas.0503272102
64 Gupta, A., Venkataraman, B., Vasudevan, M. and Gopinath Bankar, K. (2017) Co-expression network analysis of toxin-antitoxin loci in *Mycobacterium tuberculosis* reveals key modulators of cellular stress. *Sci. Rep.* **7**, 5868 https://doi.org/10.1038/s41598-017-06003-7
65 Griffin, J.E., Gawronski, J.D., Dejesus, M.A., Ioerger, T.R., Akerley, B.J. and Sassetti, C.M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**, e1002251 https://doi.org/10.1371/journal.ppat.1002251
66 Koshland, D.E. and Hamadani, K. (2002) Proteomics and models for enzyme cooperativity. *J. Biol. Chem.* **277**, 46841–46844 https://doi.org/10.1074/jbc.R200014200
67 Levitzki, A. and Koshland, D.E. (1969) Negative cooperativity in regulatory enzymes. *Proc. Natl Acad. Sci. U.S.A.* **62**, 1121–1128 https://doi.org/10.1073/pnas.62.4.1121
68 De Meyts, P. and Roth, J. (1975) Cooperativity in ligand binding: a new graphic analysis. *Biochem. Biophys. Res. Commun.* **66**, 1118–1126 https://doi.org/10.1016/0006-291X(75)90473-8
69 Ha, S.H. and Ferrell, Jr, J.E. (2016) Thresholds and ultrasensitivity from negative cooperativity. *Science* **352**, 990–993 https://doi.org/10.1126/science.aad5937
70 Overgaard, M., Borch, J. and Gerdes, K. (2009) RelB and RelE of *Escherichia coli* form a tight complex that represses transcription via the ribbon–helix–helix motif in RelB. *J. Mol. Biol.* **394**, 183–196 https://doi.org/10.1016/j.jmb.2009.09.006
71 Cataudella, I., Trusina, A., Sneppen, K., Gerdes, K. and Mitarai, N. (2012) Conditional cooperativity in toxin–antitoxin regulation prevents random toxin activation and promotes fast translational recovery. *Nucleic Acids Res.* **40**, 6424–6434 https://doi.org/10.1093/nar/gks297
72 Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R.D. et al. (2010) Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* **142**, 101–111 https://doi.org/10.1016/j.cell.2010.05.039

BIOCHEMISTRY

A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems

Yiming Cai[1]*, Ben Usher[2]*, Claude Gutierrez[3], Anastasia Tolcan[4], Moise Mansour[1], Peter C. Fineran[5,6,7], Ciarán Condon[4], Olivier Neyrolles[3], Pierre Genevaux[1†‡], Tim R. Blower[2†‡]

Toxin-antitoxin systems are widespread stress-responsive elements, many of whose functions remain largely unknown. Here, we characterize the four DUF1814-family nucleotidyltransferase-like toxins ($MenT_{1–4}$) encoded by the human pathogen *Mycobacterium tuberculosis*. Toxin $MenT_3$ inhibited growth of *M. tuberculosis* when not antagonized by its cognate antitoxin, $MenA_3$. We solved the structures of toxins $MenT_3$ and $MenT_4$ to 1.6 and 1.2 Å resolution, respectively, and identified the biochemical activity and target of $MenT_3$. $MenT_3$ blocked in vitro protein expression and prevented tRNA charging in vivo. $MenT_3$ added pyrimidines (C or U) to the 3′-CCA acceptor stems of uncharged tRNAs and exhibited strong substrate specificity in vitro, preferentially targeting $tRNA^{Ser}$ from among the 45 *M. tuberculosis* tRNAs. Our study identifies a previously unknown mechanism that expands the range of enzymatic activities used by bacterial toxins, uncovering a new way to block protein synthesis and potentially treat tuberculosis and other infections.

INTRODUCTION

Toxin-antitoxin (TA) systems are widely distributed throughout prokaryotic genomes and have been shown to help bacteria to survive predation by bacteriophages, immune responses, and antibiotic treatments (*1*–*5*). In many cases, however, the roles of chromosomal TA systems remain largely unknown, primarily due to the lack of a phenotype associated with deletion mutants under in vitro laboratory conditions (*6*–*9*). TA systems are also widespread among mobile genetic elements, including plasmids, superintegrons, cryptic prophages, and conjugative transposons, where they contribute to their stability (*10*, *11*).

TA systems encode two components, a toxic protein that targets an essential cellular process and an antagonistic antitoxin, which blocks toxin activity when cells are growing under favorable conditions. Although the processes that lead to toxin activation remain under debate, it has been proposed that under certain stress conditions, increased toxin transcription and synthesis may lead to activation (*8*, *12*). This, in turn, reduces growth rate, which can provide a means to survive with minimal metabolic burden until favorable conditions return (*13*).

TA systems are divided into six types according to the nature of the toxin and antitoxin (whether they are RNA or protein) and the mechanism of toxin antagonism (*3*). Type II systems, in which a protein toxin is sequestered by a protein antitoxin, have been most extensively studied. They are also remarkably abundant in *Mycobacterium tuberculosis*, which potentially encodes more than 80 type II TA systems, and are thought to have contributed to the success of *M. tuberculosis* as a human pathogen (*14*–*16*). Many of the putative *M. tuberculosis* toxins tested thus far were shown to inhibit bacterial growth, suggesting that these TA systems are functionally active and could modulate *M. tuberculosis* growth under certain conditions, thereby contributing to survival in the human host (*15*, *17*). Accordingly, many *M. tuberculosis* TA operons were shown to be induced in response to relevant stressors, including hypoxia, the presence of antimicrobial drugs, or macrophage engulfment (*14*, *17*). As *M. tuberculosis* encodes, among others, more than 50 VapBC, 10 MazEF, 3 HigBA, and 3 RelBE TA systems, it might be expected that there is redundancy between them, alongside condition-specific applications for each system. Furthermore, the highly toxic nature of some of these toxins suggests that their antibacterial mechanisms could be developed into antimicrobials (*18*).

This study focuses on a family of four putative toxins from *M. tuberculosis*, namely, Rv0078A, Rv0836c, Rv1045, and Rv2826c, which share a conserved nucleotidyltransferase (NTase)–like domain annotated as domain of unknown function (DUF) 1814 (Fig. 1A). The most well-characterized example of this DUF1814 family is AbiEii from *Streptococcus agalactiae*, which shares 18.3% sequence identity with Rv1045, and was identified within the AbiE abortive infection bacteriophage-defense systems (*19*). AbiEii was shown to constitute a new type of TA system, type IV, based on the observation that no interaction could be detected between the toxin and the antitoxin proteins (*20*). The DUF1814 family of proteins is widespread in bacterial, archaeal, and fungal genomes (*20*), though not all examples are genetically linked to putative antitoxins. As putative NTases, DUF1814 proteins contain four conserved motifs. The N-terminal motifs I and II are found in DNA polymerase β and are proposed to coordinate a metal ion for nucleotide binding and

[1]Laboratoire de Microbiologie et Génétique Moléculaires, Centre de Biologie Intégrative, Université de Toulouse, CNRS, UPS, 118 route de Narbonne, 31400 Toulouse, France. [2]Department of Biosciences, Durham University, South Road, Durham DH1 3LE, UK. [3]Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 205 route de Narbonne, 31400 Toulouse, France. [4]UMR8261 (CNRS, Université de Paris), Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France. [5]Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin 9054, New Zealand. [6]Genetics Otago, University of Otago, PO Box 56, Dunedin 9054, New Zealand. [7]Bio-protection Research Centre, University of Otago, PO Box 56, Dunedin 9054, New Zealand.
*These authors contributed equally to this work.
†Co–senior authors.
‡Corresponding author. Email: genevaux@ibcg.biotoul.fr (P.G.), timothy.blower@durham.ac.uk (T.R.B.)

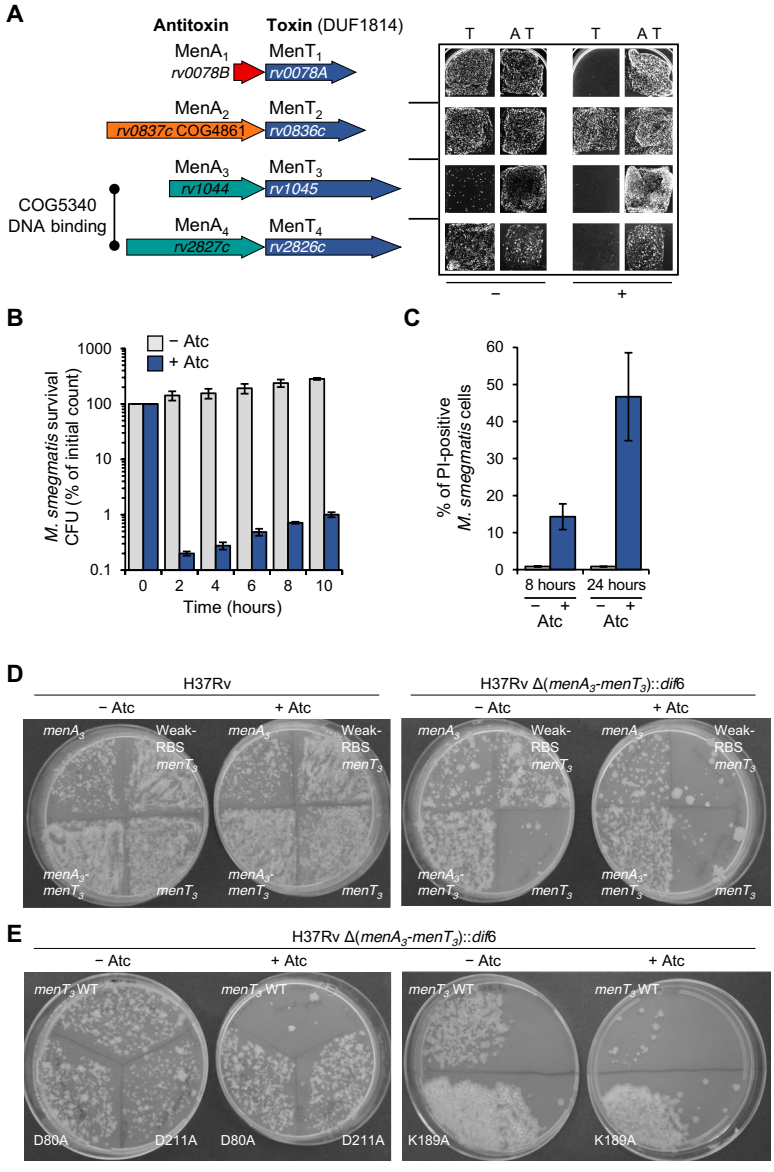


Fig. 1. Analysis of the four TA systems with NTase-like toxins encoded by the *M. tuberculosis* genome. (**A**) Scaled representation of the four *M. tuberculosis* TA systems containing NTase-like toxin genes with original and revised nomenclature (left), and corresponding toxicity and antitoxicity assays in *M. smegmatis* (right). For toxicity and antitoxicity assays, cotransformants of *M. smegmatis* mc^2 155 containing pGMC-vector, -$MenT_1$, -$MenT_2$, -$MenT_3$, or -$MenT_4$ (toxins) and pLAM-vector, -$MenA_1$, -$MenA_2$, -$MenA_3$, or -$MenA_4$ (antitoxins) were plated on LB-agar in the presence or absence of anhydrotetracycline (Atc; 100 ng ml^{-1}) and acetamide (Ace; 0.2%) inducers for toxin and antitoxin expression, respectively. Plates were incubated for 3 days at 37°C. "T" and "A" denote toxin and antitoxin, respectively. "–" and "+" represent absence or presence of inducer, respectively. (**B**) *M. smegmatis* strain mc^2 155 transformed with plasmid pGMCS-TetR-P1-RBS1-$MenT_3$ was grown in complete 7H9 medium with Sm. At time 0, the culture was divided into two. Half was kept in the same medium (pale blue bars) and half was additionally treated with Atc (200 ng ml^{-1}) (dark blue bars). Samples were harvested at the indicated times, washed, diluted, and plated on LB-agar with Sm but without Atc. Colonies were counted after 3 days at 37°C. Shown values are the average of three biological replicates with SD. CFU, colony-forming unit. (**C**) Samples of the same cultures as in (B) were harvested after 8 or 24 hours, labeled with the LIVE/DEAD BacLite dyes [Syto 9; propidium iodide (PI)], and analyzed by fluorescence-activated cell sorting. The percentage of PI-positive cells is shown for each sample (pale blue bars, no Atc; dark blue bars, 200 ng ml^{-1} Atc). Shown values are the average of three biological replicates with SD. (**D**) *M. tuberculosis* wild-type (WT) H37Rv or mutant strain H37Rv Δ(*$menA_3$-$menT_3$*)::*dif*6 were transformed with 100 ng of plasmids expressing either *$menA_3$*, *$menT_3$*, or *$menA_3$-$menT_3$*. These plasmids encode a consensus Shine-Dalgarno sequence (RBS1), except for "Weak-RBS-*$menT_3$*," which encodes a near-consensus sequence (RBS4) to weaken expression. After phenotypic expression, half of the transformation mix was plated on 7H11 oleic acid–albumin-dextrose-catalase (OADC) plates with Sm, and the other half was plated on 7H11 OADC Sm plates supplemented with Atc (200 ng ml^{-1}). Plates were imaged after 20 days at 37°C; data are representative of three independent experiments. (**E**) Mutant strain H37Rv Δ(*$menA_3$-$menT_3$*)::*dif*6 was transformed with 100 ng of plasmids expressing either *$menT_3$* WT or mutant alleles introducing the D80A, K189A, or D211A substitutions. After phenotypic expression, half of the transformation mix was plated on 7H11 OADC plates with Sm, and the other half was plated on 7H11 OADC Sm plates supplemented with Atc (200 ng ml^{-1}). Pictures were taken after 20 days at 37°C; data are representative of three independent experiments.

transfer (*20*). The C-terminal motif III is similar to that of tRNA NTases that add the 3′-CCA motif to immature tRNAs and may be important for base stacking with substrates (*21*). The C-terminal motif IV is unique to DUF1814 proteins and is proposed to form a catalytic site with motif III (*20*).

In *M. tuberculosis*, the DUF1814 toxins are encoded downstream of a variety of putative antitoxins (Fig. 1A). The toxin gene *rv0078A* is paired with a short upstream open reading frame encoding a 68–amino acid antitoxin, Rv0078B, related to MazE antitoxins, which is predicted to be disordered and lacking a DNA-binding domain (*16*). Toxin gene *rv0836c* lies downstream of a COG4861 gene, encoding a much larger putative antitoxin than the cognate toxin (Fig. 1A). Rv1045 and Rv2826c toxins are downstream of their cognate putative antitoxins Rv1044 and Rv2827c, respectively, both of which are COG5340 transcriptional regulator family proteins (Fig. 1A). COG5340 proteins include the *S. agalactiae* AbiEi antitoxin partner of AbiEii, which has previously been shown to bind to and repress the *abiE* promoter, similar to autoregulation observed in type II TA loci (*22*). An earlier transposon site hybridization study identified both the Rv1044 and Rv2827c antitoxins as essential for growth (*23*). Saturating transposon mutagenesis has additionally demonstrated that Rv1044 is essential, while transposon insertions in Rv2827c impart a growth defect (*24*). The fact that both antitoxins are important for *M. tuberculosis* growth strongly suggests that their putative cognate Rv1045 and Rv2826c toxins inhibit growth in *M. tuberculosis*.

Here, we undertook a series of microbiological, structural, genetic, and biochemical studies to investigate the DUF1814 toxins of *M. tuberculosis* and reveal their mode of action. We show that the Rv1045 toxin is a tRNA NTase that is active in *M. tuberculosis* and blocks translation through a previously undescribed mechanism involving inactivation of serine tRNAs.

RESULTS

Three DUF1814 proteins are part of bona fide TA systems

We first investigated the activity of the putative TA systems containing NTase-like DUF1814 toxins in *Mycobacterium smegmatis*, which is closely related to *M. tuberculosis* and does not encode similar antitoxins (*15*). On the basis of the findings presented below, we renamed these putative systems as "mycobacterial AbiE-like NTase toxins" (MenT) and antitoxins (MenA), numbered according to their order in the *M. tuberculosis* genome (Fig. 1A). Toxins and antitoxins were expressed in trans, with the toxins cloned into the pGMC-integrative plasmid under the control of an anhydrotetracycline (Atc)–inducible promoter and the antitoxins into the compatible pLAM plasmid under the control of an acetamide (Ace)–inducible promoter (Fig. 1A). Among the four putative toxins, only $MenT_1$ has been tested so far and was shown to be toxic in *M. smegmatis* when expressed without the upstream open reading frame encoding $MenA_1$, suggesting that $MenA_1$-$MenT_1$ form a functional TA system (*16*). Accordingly, the data presented in Fig. 1A show that $MenT_1$ toxicity was efficiently counteracted by $MenA_1$ expressed in trans. Both $MenA_3$-$MenT_3$ and $MenA_4$-$MenT_4$ also acted as TA pairs, while $MenT_2$ expression was not toxic (Fig. 1A). Inhibition of $MenT_4$ toxicity could only be achieved when the putative antitoxin was expressed in the context of the *menA*$_4$-*menT*$_4$ operon (Fig. 1A). Expression of $MenA_4$ alone from pLAM was toxic (fig. S1A), indicating that $MenA_4$-$MenT_4$ might not function as a typical TA pair under these conditions. Similar experiments performed in *Escherichia coli* confirmed the phenotypes observed in *M. smegmatis* for $MenA_2$-$MenT_2$, $MenA_3$-$MenT_3$, and $MenA_4$-$MenT_4$ (including $MenA_4$ toxicity), but not for $MenT_1$, which exhibited no detectable toxicity in *E. coli* (fig. S1B). Last, coexpression of the active toxins with noncognate antitoxins did not reveal any detectable cross-talk between the different TA pairs (fig. S1, A and C). Note that cross-talk assays with $MenA_4$ antitoxin expressed from pLAM in *M. smegmatis* could not be performed because of its toxicity.

Ectopic expression of $MenT_3$ in the presence of inducer showed the most robust toxicity in both *M. smegmatis* and *E. coli* when compared to the other toxins (Fig. 1A and fig. S1). In *M. smegmatis*, only a few $MenT_3$ transformants were obtained, even in the absence of inducer. Ectopic expression of $MenT_3$ in *M. smegmatis* induced a rapid drop of about 3-$\log_{10}$ in colony-forming units only 2 hours after induction with Atc (Fig. 1B). LIVE/DEAD BacLight stains have previously been used to study the effects of toxin expression on cell viability in *M. tuberculosis* (*18*). Flow cytometry analysis of *M. smegmatis* expressing $MenT_3$ revealed that the proportion of propidium iodide–permeable cells was substantially higher in $MenT_3$-induced versus noninduced cells 8 or 24 hours after induction with Atc (Fig. 1C), indicating that $MenT_3$ strongly affects cell viability.

$MenA_3$-$MenT_3$ is a functional TA system in *M. tuberculosis*

To investigate the impact of $MenA_3$ and $MenT_3$ on *M. tuberculosis* growth, plasmids encoding the toxin, the antitoxin, or both, were introduced into H37Rv wild-type (WT) strain. The resulting transformants were not sensitive to ectopic expression of $MenT_3$ (Fig. 1D), presumably because endogenous $MenA_3$ was sufficient to neutralize the sum of endogenous and ectopic $MenT_3$. To confirm this hypothesis, we attempted to construct a strain deleted for the *menA*$_3$-*menT*$_3$ operon. Previous work showed that *menA*$_3$ cannot be disrupted by transposon insertion (*24*). Accordingly, we found that deletion of the *menA*$_3$-*menT*$_3$ operon in *M. tuberculosis* H37Rv strain could not be achieved, most likely because simultaneous disruption of both genes resulted in a toxic effect from residual $MenT_3$. To circumvent this problem, we constructed the deletion in a derivative of H37Rv carrying a second copy of *menA*$_3$ constitutively expressed from a pGMC integrative plasmid. Once the *menA*$_3$-*menT*$_3$ operon was deleted, it was then possible to remove the ectopic copy of *menA*$_3$ by pGMC plasmid replacement (fig. S2). The Δ*menA*$_3$-*menT*$_3$ mutant became highly sensitive to the $MenT_3$ toxin, even in the absence of inducer (Fig. 1D). Therefore, to finally obtain transformants, *menT*$_3$ was cloned downstream of a weaker Shine-Dalgarno sequence. Using this construct, we observed inducible $MenT_3$ toxicity, which was fully abolished by the presence of the antitoxin (Fig. 1D). Together, these data demonstrate that the $MenT_3$ toxin inhibits growth and that $MenA_3$-$MenT_3$ functions as a bona fide TA pair in *M. tuberculosis*.

A previous amino acid sequence alignment of DUF1814 putative NTases highlighted conserved residues, a number of which were confirmed as essential for AbiEii toxicity in *S. agalactiae* (*20*). To investigate whether some of these residues were important for $MenT_3$ toxicity, we selected and engineered three conserved residues for substitution: D80A, localized in the DNA polβ superfamily motif, and K189A and D211A, both toxin-specific residues. We then tested the impact of these substitutions on *M. tuberculosis* growth (Fig. 1E). All three substitutions abolished $MenT_3$ toxicity in both *M. tuberculosis* (Fig. 1E) and *E. coli* (fig. S3A).

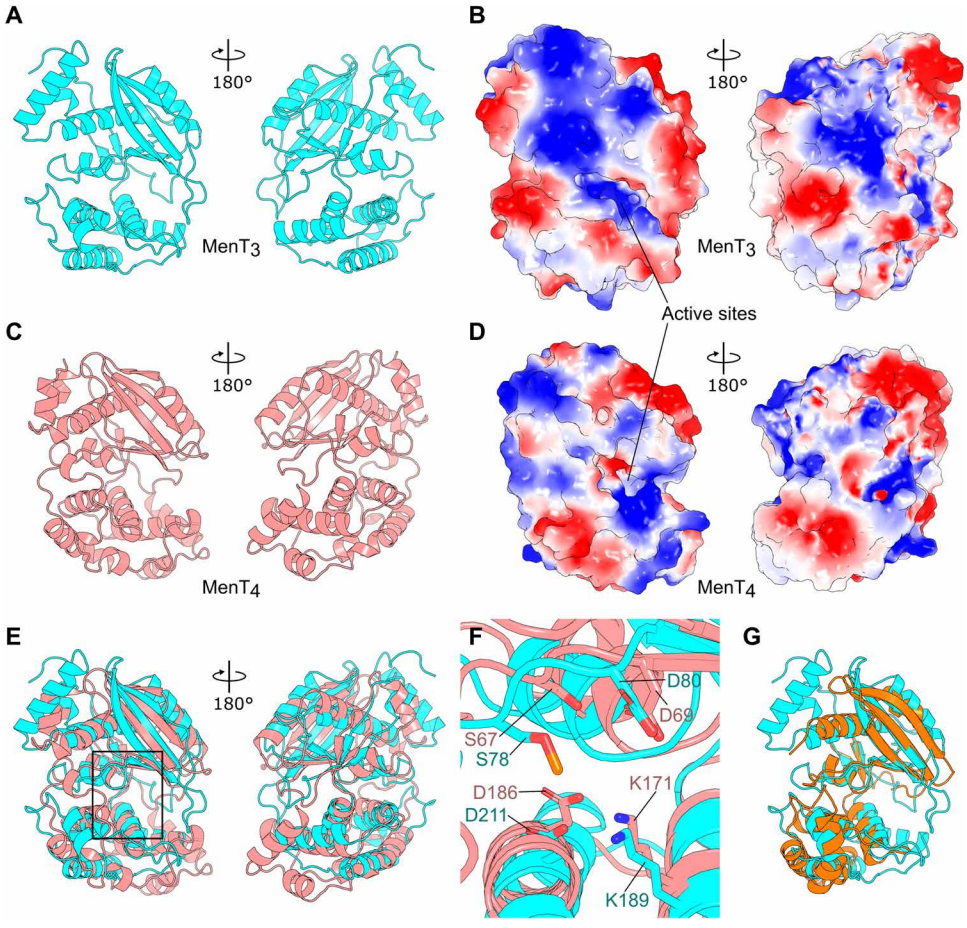


Fig. 2. Crystal structure of the MenT$_3$ and MenT$_4$ toxins. (**A**) Structure of monomeric MenT$_3$ toxin, with views from front and back, shown as cyan cartoon representations. (**B**) Surface electrostatics of MenT$_3$, viewed as in (A), with red for electronegative and blue for electropositive potential. (**C**) Structure of monomeric MenT$_4$, with views from front and back, shown as salmon cartoon representations. (**D**) Surface electrostatics of MenT$_4$, viewed as in (C), colored as per (B). (**E**) Superposition of MenT$_4$ onto MenT$_3$, viewed and colored as per (A) and (C). (**F**) Tilted close-up view of the toxin active sites, as indicated by the boxed region of (E). MenT$_3$ residues S78 (phosphorylated), D80, K189, and D211 are indicated, along with the homologous MenT$_4$ residues S67, D69, K171, and D186. (**G**) Alignment of JHP933 (PDB: 4O8S) as orange cartoon representation, against MenT$_3$ viewed and colored as per (A, left).

Next, we investigated whether the MenT$_3$ toxin and MenA$_3$ antitoxin could interact in vivo. Since this TA pair is functional in *E. coli*, we performed affinity-tagged in vivo copurification experiments in *E. coli* using His-tagged variants of MenT$_3$ and MenA$_3$, which were first confirmed to be active as toxin and antitoxin, respectively (fig. S4A). In strains coexpressing both the toxin and the antitoxin (with either the toxin or the antitoxin tagged), and with tagged toxin and tagged antitoxin alone as controls, the in vivo copurification revealed that a small but significant fraction of the MenT$_3$ toxin and the MenA$_3$ antitoxin copurified, whether the toxin or the antitoxin was used as bait (fig. S4, C and D). Similar results were obtained with the MenA$_1$-MenT$_1$ pair, which encodes a much shorter, unrelated antitoxin (fig. S4, B, E, and F). Together, these data show that both TA partners can interact, but it remains to be determined whether a direct interaction between an NTase toxin and its cognate antitoxin is required for toxin inhibition.

MenT$_3$ and MenT$_4$ are structural homologs

To begin investigations into the mechanism of toxicity of MenT$_3$, we solved its structure to 1.6 Å resolution by x-ray crystallography (Fig. 2A and Table 1). MenT$_3$ is a monomeric bi-lobed globular protein, with two hemispheres connected by a short linker (Fig. 2A). This monomeric assembly matches the expected size observed by size exclusion chromatography. Surface electrostatics show a distinct electropositive surface leading to a deeper recess (Fig. 2B, left), which contains residues D80, K189, and D211 that were needed for toxicity in vivo (Fig. 1E). This potentially indicates the position of the active site, and the electropositive surface may facilitate interaction with electronegative substrates such as nucleic acids. To further characterize the DUF1814 family, we also solved the MenT$_4$ toxin structure to 1.2 Å resolution (Fig. 2C and Table 1). MenT$_4$ is also monomeric (also observed by size exclusion chromatography) and the overall architecture is similar to, but not exactly the same as, MenT$_3$. MenT$_4$ has a bi-lobed globular structure and distinct electropositive patches close to a similarly positioned active site region (Fig. 2, C and D). Aligning MenT$_3$ and MenT$_4$ by sequence gave a poor root mean square deviation (RMSD) of 13.4 Å; however, this can be improved to 4.7 Å using sequence-independent superposition, which demonstrates similarity in overall fold (Fig. 2E). A close-up of MenT$_3$ residues D80, K189, and D211 show them

Table 1. Crystallographic data collection and refinement statistics.

	$MenT_3$ native	$MenT_3$ Se-peak	$MenT_3$ Se-high remote	$MenT_3$ Se-inflection	$MenT_4$ native
Data collection					
PDB ID code	6Y5U	-	-	-	6Y56
Beamline	Diamond I04	Diamond I03	Diamond I03	Diamond I03	Diamond I24
Wavelength (Å)	0.9795	0.9793	0.9641	0.9795	0.9781
Resolution range (Å)	47.70–1.59 (1.65–1.59)*	47.78–2.19 (2.26–2.19)	47.83–2.05 (2.11–2.05)	53.13–2.04 (2.11–2.04)	42.23–1.23 (1.27–1.23)
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P2_1$
Unit cell, *a b c* (Å), α β γ (°)	95.4 95.4 69.0, 90.0 90.0 120.0	95.6 95.6 69.2, 90.0 90.0 120.0	95.7 95.7 69.3, 90.0 90.0 120.0	95.6 95.6 69.3, 90.0 90.0 120.0	42.3 57.8 54.7, 90.0 92.3 90.0
Total reflections	98,016 (9668)	36,407 (3179)	44,514 (3476)	47,255 (4637)	149,653
Unique reflections	49,008 (4834)	19,130 (1646)	23,313 (1788)	23,628 (2319)	75,996 (7206)
Multiplicity	2.0	1.9	1.9	2.0	2.0
Completeness (%)	99.95 (99.83)	100.00 (100.00)	100.00 (99.80)	100.00 (99.70)	98.80 (88.97)
Mean $I/\sigma(I)$	16.7	6.5	7.6	8.9	7.0
R_{merge}	0.016 (0.486)	0.055 (0.373)	0.055 (0.522)	0.048 (0.463)	0.060 (0.926)
R_{meas}	0.022 (0.687)	0.077 (0.528)	0.078 (0.739)	0.068 (0.654)	0.085 (1.310)
$CC_{1/2}$	1.0 (0.672)	0.995 (0.803)	0.997 (0.544)	0.997 (0.641)	0.996 (0.294)
Refinement					
R_{work}	0.2024 (0.2924)	-	-	-	0.1840 (0.3174)
R_{free}	0.2242 (0.3108)	-	-	-	0.1950 (0.3352)
No. of non-hydrogen atoms	2494	-	-	-	2649
Macromolecules	2213	-	-	-	2322
Solvent	281	-	-	-	327
Protein residues	288	-	-	-	292
RMSD (bonds, Å)	0.006	-	-	-	0.005
RMSD (angles, °)	0.940	-	-	-	0.830
Ramachandran favored (%)	97.53	-	-	-	98.28
Ramachandran allowed (%)	2.47	-	-	-	1.72
Ramachandran outliers (%)	0.00	-	-	-	0.00
Average *B* factor	34.1	-	-	-	20.6
Macromolecules	33.4	-	-	-	19.3
Solvent	39.4	-	-	-	29.9

*Statistics for the highest-resolution shell are shown in parentheses.

clustered at the putative active site, and when overlaid, the homologous $MenT_4$ residues, D69, K171, and D186, respectively, take up similar positions (Fig. 2F). There was also density for a phosphoserine at $MenT_3$ S78, but the corresponding residue in $MenT_4$, S67, was not phosphorylated (Fig. 2F). Searches for structural homologs of $MenT_3$ and $MenT_4$ were performed using the DALI server (*25*). Among multiple hits for NTases, the best match was for JHP933 from *Helicobacter pylori*, a predicted NTase encoded by the *jhp0933* gene (*26*). JHP933 aligned to $MenT_3$ with an RMSD of 2.4 Å, though multiple additional helices were resolved in the $MenT_3$ structure (Fig. 2G). An analysis of the *H. pylori* genome revealed that the *jhp0932* gene lies just upstream of *jhp0933* and partially overlaps its coding sequence. The presence of these genes in what appears to be a classic TA configuration suggests that JHP933 may belong to the $MenT_3$/$MenT_4$ family of NTase-like toxins.

RNase PH overexpression confers resistance to $MenT_3$

$MenT_3$ is the most toxic of the four *M. tuberculosis* NTase-like toxins tested, both in mycobacteria and in *E. coli* (Fig. 1 and fig. S1). We therefore took advantage of this robust toxicity to search for *E. coli* genes that were able to suppress $MenT_3$-mediated growth inhibition when overexpressed. We reasoned that identification of such

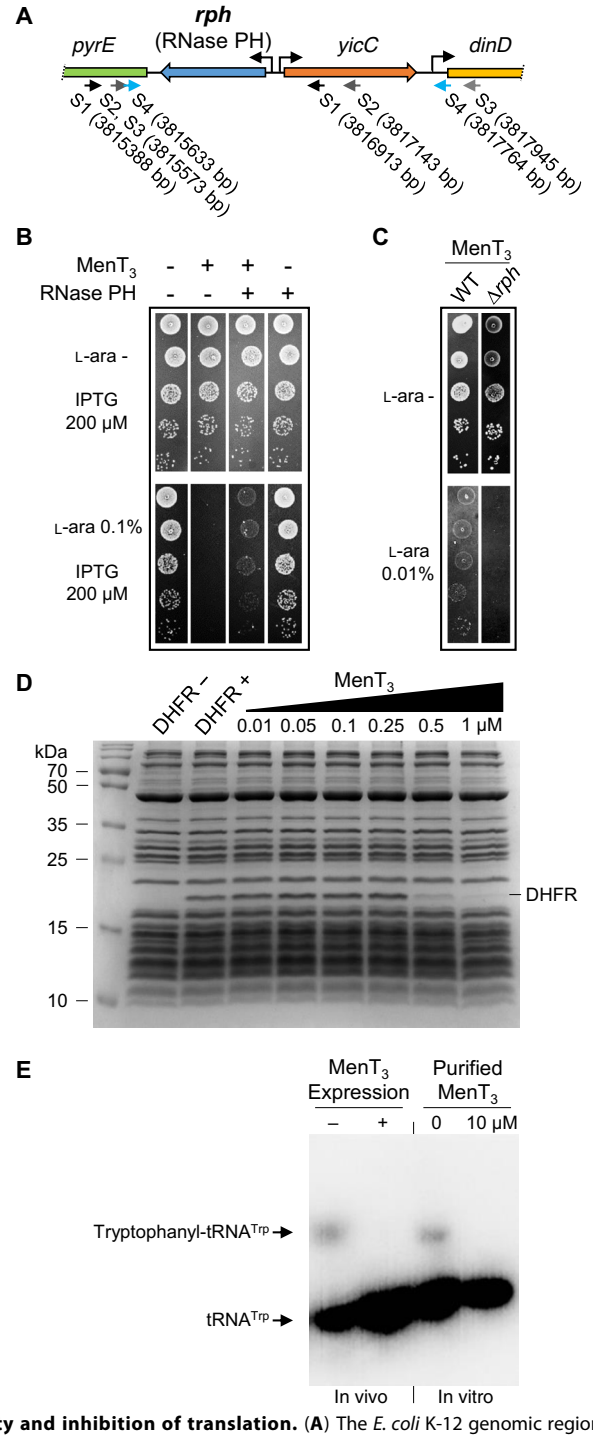


Fig. 3. RNase PH suppresses MenT$_3$ toxicity and inhibition of translation. (**A**) The *E. coli* K-12 genomic region containing the *rph* gene is shown. Suppressor plasmids that counteract MenT$_3$ toxicity encoded *rph*, as depicted by small arrows under the adjacent genes *pyrE*, *yicC*, and *dinD*. The positions in base pair of the ends of each suppressor fragment, in relation to the *E. coli* K-12 chromosome, are indicated between brackets. (**B**) Overexpression of *E. coli* RNase PH partially suppresses MenT$_3$ toxicity. *E. coli* DLT1900 strains containing either pK6-vector (–) or pK6-MenT$_3$ (+) were cotransformed with p29SEN-vector (–) or p29SEN-Rph (RNase PH) (+). The resulting cotransformants were serially diluted, spotted onto LB-agar plates in the presence or absence of L-ara (0.1%) and IPTG (200 μM) inducers, and incubated at 37°C. (**C**) Deletion of *rph* further increases MenT$_3$ toxicity. Transformants of *E. coli* DLT1900 WT and Δ*rph* mutant strains containing plasmid pK6-MenT$_3$ were serially diluted, spotted onto LB-agar plates with or without L-ara (0.01%), and incubated at 37°C. (**D**) In vitro transcription/translation reactions assessing levels of DHFR control protein produced in the absence or presence of increasing concentrations of MenT$_3$ toxin. Samples were separated by SDS–polyacrylamide gel electrophoresis and stained with InstantBlue. (**E**) For in vivo assays, transformants of *E. coli* BL21 (λDE3) containing plasmid pET-MenT$_3$ or the empty vector were grown in M9M at 37°C. Following overexpression of MenT$_3$, tRNAs were extracted, separated, and visualized by Northern blot using specific radiolabeled probes against tRNA^{Trp}. For in vitro assays, purified MenT$_3$ (10 μM) was added to transcription/translation assays producing GatZ protein. After 2 hours at 37°C, tRNAs were extracted, separated, and visualized by Northern blot as performed for the in vivo samples. All images are representative of triplicate data.

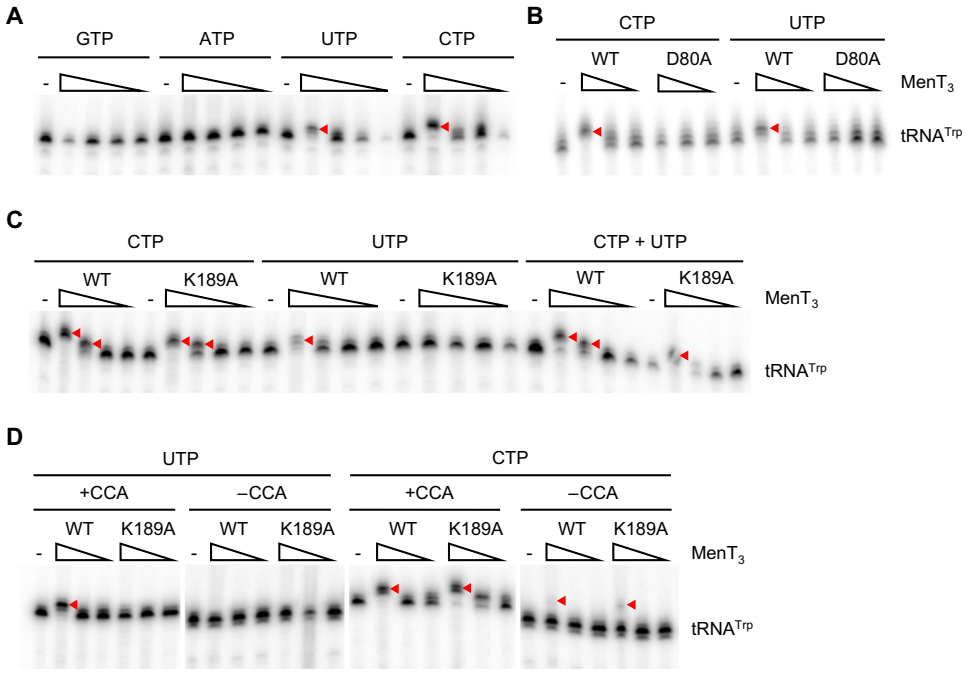


Fig. 4. Toxin $MenT_3$ adds pyrimidines to the 3′-CCA acceptor stem of tRNA. (**A**) Radiolabeled *E. coli* $tRNA^{Trp}$ was incubated with 1, 0.1, 0.01, or 0.001 μg of $MenT_3$ WT or no toxin (–) for 20 min at 37°C in the presence of unlabeled GTP, ATP, UTP, or CTP. Extended products are indicated with arrowheads throughout all panels. (**B**) Radiolabeled *E. coli* $tRNA^{Trp}$ was incubated with 1, 0.1, or 0.01 μg of $MenT_3$ WT or $MenT_3^{(D80A)}$ with CTP or UTP, as per conditions in (A). (**C**) Incubation of radiolabeled *E. coli* $tRNA^{Trp}$ with 1, 0.1, 0.01, or 0.001 μg of $MenT_3$ WT or $MenT_3^{(K189A)}$, with CTP, UTP, or a mixture of both, as per conditions in (A). (**D**) Radiolabeled *E. coli* $tRNA^{Trp}$ preparations, made with or without a 3′-CCA motif, were incubated with 1, 0.1, or 0.01 μg of either $MenT_3$ WT, $MenT_3^{(K189A)}$, or no toxin (–), for 20 min at 37°C in the presence of unlabeled UTP or CTP. Note that the (–) CCA lanes have been overexposed to equalize intensity to the (+) CCA lanes of the same gel. Assays of the individual WT and $MenT_3$ substitution proteins and $tRNA^{Trp}$ ± CCA substrates shown in (A) to (D) were performed between two and four times.

suppressors might potentially shed light on the cellular processes affected by the toxin. Details of the genetic selection used are described in Materials and Methods. Among the approximately 60,000 clones of the *E. coli* genomic plasmid library tested in this work, we identified 18 plasmids that passed two rounds of selection and appeared to encode bona fide suppressors of $MenT_3$ toxicity. We observed that the toxin-resistant colonies were noticeably smaller and translucent compared to noninduced cells, indicating that, although notably reduced, $MenT_3$ toxicity is not fully suppressed. Sequencing of the genomic regions encoded by the 18 suppressor plasmids revealed that several of these candidate plasmids harbored the same genomic fragments. Six different suppressor clones encompassing two different regions of the *E. coli* chromosome were identified. Two of the six suppressor plasmids harbored the *ydeA* gene, encoding an L-arabinose (L-ara) exporter protein known to decrease L-ara levels in *E. coli* (*27*). These suppressors were discarded as YdeA overexpression would presumably decrease toxicity of many toxic proteins expressed from the *araBAD* promoter. The four other suppressor plasmids harbored the *rph* gene, encoding the phosphorolytic ribonuclease (RNase PH), involved in the 3′ processing of RNAs (Fig. 3A). RNase PH removes nucleotides downstream of the 3′-CCA sequence, required for aminoacylation of tRNAs, from tRNA precursors with 3′ extensions. It is also involved in other RNA maturation and quality control processes, including the maturation of rRNA (*28*).

Suppression of $MenT_3$ toxicity by RNase PH overexpression was confirmed by cloning *rph* alone in a low–copy number plasmid under the control of an isopropyl-β-D-thiogalactopyranoside (IPTG)–inducible promoter and assaying for growth in the presence of $MenT_3$ in *E. coli* (Fig. 3B). We also showed that the toxicity of $MenT_3$ was enhanced when expressed in *E. coli* carrying a deletion of the *rph* gene, even with a 10-fold decrease in inducer levels (Fig. 3C), further reinforcing the genetic link between *menT*$_3$ and *rph*. The primary role of RNase PH in processing tRNAs suggests that DUF1814 NTase-like toxins could act directly at the site of aminoacylation at the 3′-end of tRNA, thus inhibiting translation. Whether endogenous RNase PH would be sufficiently induced in response to toxin expression to help restore the functional tRNA pool in recovering *M. tuberculosis* cells remains to be determined.

$MenT_3$ inhibits tRNA charging

$MenT_3$ WT and the $MenT_3^{(D80A)}$ and $MenT_3^{(K189A)}$ substitutions were overexpressed and purified for biochemical characterization. When tested in an in vitro transcription/translation reaction that uses recombinant *E. coli* components, purified $MenT_3$ WT reduced production of the *E. coli* dihydrofolate reductase (DHFR) control protein in a concentration-dependent manner (Fig. 3D). Compared to $MenT_3$ WT, $MenT_3^{(D80A)}$ and $MenT_3^{(K189A)}$ had a markedly reduced impact on the production of DHFR (fig. S5A). The same trend was observed when $MenT_3$ WT, $MenT_3^{(D80A)}$, and $MenT_3^{(K189A)}$ were used in in vitro reactions producing WaaF and GatZ as test proteins (fig. S5, B and C). We also expressed and purified $MenT_4$ WT and demonstrated that this, too, prevented the production of DHFR in a concentration-dependent manner in in vitro transcription/translation assays (fig. S5D).

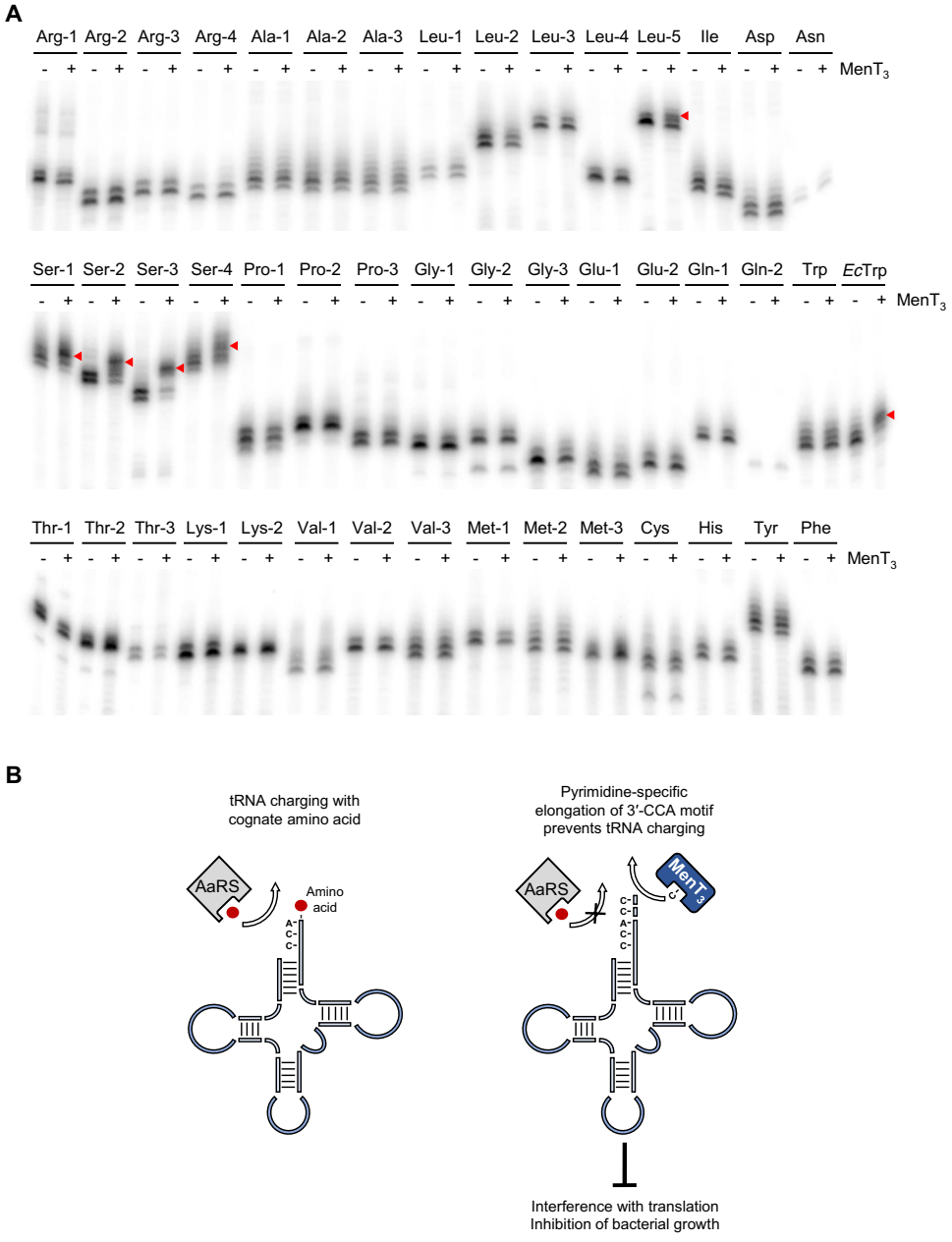


Fig. 5. Screening for $MenT_3$ *M. tuberculosis* tRNA targets. (**A**) Radiolabeled *M. tuberculosis* tRNAs were incubated with 0.1 μg of $MenT_3$ WT (+) or no toxin (–) for 20 min at 37°C in the presence of unlabeled CTP. *E. coli* $tRNA^{Trp}$ (EcTrp) was used as a positive control. The global screen of all *M. tuberculosis* tRNA was performed once and the effect of $MenT_3$ $tRNA^{Ser2}$ was confirmed twice independently. (**B**) Schematic diagram of the $MenT_3$ toxin mechanism of action. $MenT_3$ elongates the 3′-CCA motif of specific tRNAs, preventing their charging by aminoacyl-tRNA synthetases (AaRS), thereby interfering with translation and inhibiting bacterial growth.

The fact that $MenT_3$ inhibited protein synthesis, and that RNase PH is involved in the removal of nucleotides following the 3′-CCA sequence required for tRNA aminoacylation, suggested that tRNA charging might be affected by $MenT_3$ expression in vivo. To address this hypothesis, we first used a method developed for *E. coli*, which separates charged from uncharged tRNAs and allows their detection by Northern blot after extraction in vivo (*29*). We chose $tRNA^{Trp}$ as a model tRNA because (i) the tryptophanyl-tRNA can be well separated from uncharged $tRNA^{Trp}$ and (ii) there is only one $tRNA^{Trp}$ in *E. coli* (*29*). No charged tryptophanyl-$tRNA^{Trp}$ could be detected following overexpression of $MenT_3$ when compared to the empty vector control (Fig. 3E and fig. S5E). $tRNA^{Trp}$ charging levels were also investigated in vitro by adding purified $MenT_3$ to the transcription/translation assay described above (Fig. 3E). In this case, $MenT_3$ also affected $tRNA^{Trp}$ charging in vitro, thus supporting the hypothesis that the toxin inhibits protein synthesis by preventing aminoacylation of tRNA.

$MenT_3$ transfers pyrimidines to the 3′ acceptor stem of specific tRNAs

The observation that $MenT_3$ is related to NTases (Fig. 2G) suggests that its mode of action is to directly transfer nucleotides to tRNAs, thereby preventing aminoacylation. We performed assays using

radiolabeled tRNAs to track the addition of nucleotides by $MenT_3$ WT, $MenT_3^{(D80A)}$, and $MenT_3^{(K189A)}$ (Fig. 4).

$MenT_3$ WT was incubated with $tRNA^{Trp}$ from *E. coli*, as a model recipient tRNA, in the presence of guanosine 5′-triphosphate (GTP), adenosine 5′-triphosphate (ATP), uridine 5′-triphosphate (UTP), or cytidine 5′-triphosphate (CTP), and nucleotide transfer was monitored as an increase in tRNA size by high-resolution polyacrylamide gel electrophoresis (PAGE; Fig. 4A). At high concentrations of the enzyme, we found that $MenT_3$ can add two to three extra nucleotides to $tRNA^{Trp}$ in the presence of CTP or UTP, with a slight preference for CTP, suggesting that $MenT_3$ is a pyrimidine-specific NTase (Fig. 4A). No transfer was observed with purines ATP or GTP as substrates (Fig. 4A). $MenT_3^{(D80A)}$, which was unable to inhibit in vitro protein synthesis (fig. S5, A to C), had no NTase activity with either UTP or CTP (Fig. 4B). $MenT_3^{(K189A)}$, which was also inactive in the in vitro transcription/translation assay (fig. S5, A to C), only lost its NTase activity in the presence of UTP, but retained some activity (albeit less than WT) in the presence of CTP, or both nucleotides (Fig. 4C). This could imply that K189A is important for substrate nucleotide selectivity. No synergistic effect was seen when $MenT_3$ WT was incubated with a mixture of CTP and UTP, as the pattern with both nucleotides together resembled that of CTP alone (Fig. 4C).

Canonical tRNA NTases typically add the 3′-CCA motif to tRNAs lacking an encoded 3′-CCA that are processed at the level of the discriminator nucleotide (nucleotide 73). They also repair this motif when 3′-exoribonucleases, such as RNase PH, fail to stop at the 3′-CCA motif when processing tRNA precursors containing an encoded 3′-CCA, typically removing the terminal A residue. Since *M. tuberculosis* contains a mixture of tRNA genes encoding or lacking a 3′-CCA motif, we wondered whether $MenT_3$ had a preference for one class (or another class) of substrate. While faint NTase activity was observed when $MenT_3$ WT and $MenT_3^{(K189A)}$ were incubated with CTP and $tRNA^{Trp}$ lacking a 3′-CCA, the data show that $MenT_3$ had a clear preference for tRNAs that already possessed a 3′-CCA motif (Fig. 4D). This is in contrast to the normal function of tRNA NTases, which prefer tRNAs lacking an intact 3′-CCA. Again, $MenT_3$ WT modified $tRNA^{Trp}$ using both CTP and UTP as substrate, while $MenT_3^{(K189A)}$ could only use CTP (Fig. 4D). Addition of nucleotides to mature tRNAs by $MenT_3$ would completely abolish the ability of these tRNAs to be charged with their cognate amino acid and take part in translation, accounting for their cellular toxicity.

Our in vivo data show that toxins $MenT_3$, as well as $MenT_1$ and $MenT_4$, are significantly less toxic in *E. coli* than in mycobacteria (Fig. 1 and fig. S1), which suggests that these toxins may have a tRNA target preference. We therefore asked whether $MenT_3$ would exhibit some specificity toward the different tRNAs of *M. tuberculosis*. We made polymerase chain reaction (PCR) templates allowing us to in vitro transcribe the 45 different tRNAs of *M. tuberculosis*, each with a 3′-CCA motif (fig. S6). As before, each radiolabeled tRNA was incubated with $MenT_3$ and nonradiolabeled CTP (Fig. 5A). To our surprise, $MenT_3$ appeared to be highly specific, preferentially modifying the four *M. tuberculosis* $tRNA^{Ser}$ isoacceptors, along with weak modification of $tRNA^{Leu5}$ (Fig. 5A). Although we cannot exclude that $MenT_3$ can modify other tRNAs in vivo, the data show that the toxin presents a high degree of specificity toward different tRNAs in vitro, which may explain the variable toxicity observed in different bacteria.

Last, we asked whether the antitoxin $MenA_3$ inhibited the NTase activity of $MenT_3$ directly, or whether it could simply reverse its action by removing the added nucleotides in a manner similar to the RNase PH multicopy suppressor. Addition of $MenA_3$ strongly inhibited the NTase activity of $MenT_3$ on the natural substrate *M. tuberculosis* $tRNA^{Ser2}$ when coincubated with the toxin at a molar ratio > 2.5. However, $MenA_3$ failed to remove the added nucleotides from $tRNA^{Ser2}$ when added after a preincubation of $tRNA^{Ser2}$ with $MenT_3$, even at high concentrations (fig. S7). This suggests that the antitoxin is likely to inhibit the toxin rather than reverse the reaction on the substrate.

DISCUSSION

This study has characterized a family of TA systems from *M. tuberculosis* containing NTase-like DUF1814 toxins, establishing $MenT_3$ as a potent toxin in this problematic pathogen. We have solved the structures of the homologous toxins $MenT_3$ and $MenT_4$ by x-ray crystallography, revealing fold similarity and conserved residues within the proposed active sites, and have observed a similar mode of toxin activity, targeting protein synthesis. We have further elucidated the mechanism of toxicity for $MenT_3$, showing that it functions as a pyrimidine-specific NTase preferentially targeting *M. tuberculosis* $tRNA^{Ser}$ in vitro (Fig. 5B).

The observation that the three NTase toxins identified in this work show different levels of toxicity when expressed in the same host, and that such toxic signatures can vary when expressed in different bacterial hosts (i.e., *E. coli* versus mycobacteria), is intriguing (Fig. 1 and fig. S1). The most marked example is $MenT_1$, which shows robust toxicity in *M. smegmatis* but no toxicity in *E. coli* (Fig. 1A and fig. S1B). Although we cannot exclude this being a result of improper folding or expression of the toxin in *E. coli*, it is also reasonable to assume that the toxin may not be able to recognize its tRNA targets due, for example, to tRNA modification, or the absence of its preferred tRNA target (*30*). Another possibility is that tRNA targets are expressed at higher levels in *E. coli* and are thus sufficiently abundant to overcome the noxious effect of the toxin in vivo. The fact that *M. tuberculosis* and *M. smegmatis* only have 45 and 46 tRNA genes, respectively, while *E. coli* has 86, is in line with this hypothesis (*30*, *31*).

The apparent in vitro specificity of $MenT_3$ for certain *M. tuberculosis* tRNAs, especially $tRNA^{Ser}$, is remarkable (Fig. 5A). We did not test other tRNAs in *E. coli* besides $tRNA^{Trp}$; it may well have been fortuitous that the only tRNA we tested in this organism was detectably modified by the toxin in vitro (Fig. 4A) and in vivo, inferred from the reduced charging levels following toxin expression (Fig. 3E). We checked whether the *M. tuberculosis* tRNAs that were substrates of $MenT_3$ had any distinguishing features and were struck by the fact that all serine tRNAs and several leucine tRNAs were unique among *M. tuberculosis* tRNAs in that they had long variable arms (fig. S8A) (*32*). While this is intriguing and may contribute to substrate specificity, it cannot be the only recognition element because (i) two leucine tRNAs besides $tRNA^{Leu5}$ have variable loops but are not $MenT_3$ substrates in vitro and (ii) *E. coli* $tRNA^{Trp}$ does not have a variable loop (fig. S8B), but can be extended by the NTase activity of the toxin. It is also intriguing in this regard that *M. tuberculosis* $tRNA^{Trp}$ is not a $MenT_3$ substrate in vitro. *E. coli* and *M. tuberculosis* $tRNA^{Trp}$ are highly homologous but do show differences in their variable- and T-arm sequences (fig. S8C). Substrate specificity therefore appears to come from a combination of multiple sequence and structure

motifs. Having identified these tRNA targets in vitro, further work is now needed to confirm targeting in vivo in *M. tuberculosis*.

Our observed TA interactions raise questions regarding the molecular mechanisms of antitoxicity for DUF1814 toxins (fig. S4, C to F). Typically, in type II TA systems, antitoxin function is in part driven by its strong and direct interaction with the cognate toxin (*3*). While we have shown interactions between cognate toxins and antitoxins (fig. S4, C to F), the antitoxin interaction in vivo appears weak. We additionally demonstrated that coincubation of the $MenA_3$ antitoxin with $MenT_3$ is able to neutralize the NTase activity (fig. S7). This suggests that any interaction-based antitoxicity might be a transient and labile mechanism and, due to the difference in size and sequence between antitoxins $MenA_1$ and $MenA_3$ (Fig. 1A), may well differ between these systems.

The DUF4433 DarT toxin from *M. tuberculosis* was recently identified as a single-stranded DNA NTase that specifically and reversibly adenosine 5′-diphosphate (ADP)–ribosylates thymidines (*33*). Our study identifies $MenT_3$ as an NTase toxin from the unrelated DUF1814 protein family. In comparison to DarT, $MenT_3$ acts via a distinct and novel mode of toxicity where the $MenT_3$ toxin preferentially targets *M. tuberculosis* tRNAs in vitro, preventing their charging with cognate amino acids by adding nucleotides to the 3′-CCA acceptor stem (Fig. 5B). Accordingly, antitoxin function also appears to differ between these systems. Whereas DarT is counteracted enzymatically by the cognate antitoxin DarG via target de-ADP-ribosylation (*33*), we found that $MenA_3$ was unable to reverse $MenT_3$ toxicity by removing nucleotides, suggesting that $MenA_3$ likely inhibits the toxin activity.

Increasing numbers of toxins have been identified that target tRNAs by various mechanisms (*13*). The *M. tuberculosis* type II VapC toxins function as endoribonucleases cleaving tRNAs (*34*), whereas TacT from *Salmonella* Typhimurium and AtaT from *E. coli* are tRNA acetyltransferases, modifying charged tRNAs to block translation (*35*, *36*). That $MenT_3$ provides yet another way to inhibit tRNA activity is perhaps not unusual, given the essential nature of translation to cellular growth and survival. This likely reflects the value of possessing multiple TA systems to promote adaptability to different stressful environments via tRNA metabolism, with downstream effects ranging from stalling cell growth to potentially altering translation output (*13*). It remains to be seen whether this mechanism is conserved among DUF1814-toxins; while $MenT_4$ shares structural similarities to $MenT_3$ and inhibits protein synthesis in vitro (Fig. 2 and fig. S5D), we have not yet explored the molecular mechanism behind its toxicity. Given the continued significance of *M. tuberculosis* worldwide, the mechanism used by the $MenA_3$-$MenT_3$ TA system highlights a new way to block protein synthesis. We propose that further exploring the molecular mechanisms of both toxicity and antitoxicity will provide useful insights into the regulation of bacterial growth.

MATERIALS AND METHODS

Bacterial strains and culture conditions

E. coli DH5α (Invitrogen), DH10B (Thermo Fisher Scientific), BL21 (λDE3) (Novagen), ER2566 (New England Biolabs), W3110 [strain American Type Culture Collection (ATCC) 27325], DLT1900 (*37*), and *M. smegmatis* mc^2 155 (strain ATCC 700084) are as previously described. To construct BL21 (λDE3) Δ*slyD*, the Δ*slyD*::Km^R allele from JW3311 (Keio collection) was moved into BL21 (λDE3) using bacteriophage P1-mediated transduction. To construct the unmarked DLT1900 Δ*rph* mutant, the Δ*rph*::Km^R allele from JW3618 (Keio collection) was first moved into DLT1900 using bacteriophage P1-mediated transduction and by subsequent removing of the kanamycin (Km) resistance cassette using plasmid pCP20, as previously described (*38*). *E. coli* were routinely grown at 37°C in LB medium or M9 minimal (M9M) medium supplemented when necessary with Km (50 μg ml^{-1}), ampicillin (Ap; 50 μg ml^{-1}), chloramphenicol (Cm; 34 μg ml^{-1}), streptomycin (Sm; 25 μg ml^{-1}), spectinomycin (Sp; 50 μg ml^{-1}), IPTG (1 mM), L-ara (0.1% w/v), or D-glucose (glu; 0.2% w/v). *M. smegmatis* mc^2 155 strains were routinely grown at 37°C in either LB or 7H9 medium (Difco). *M. tuberculosis* H37Rv (WT; ATCC 27294) and mutant strains were routinely grown at 37°C in complete 7H9 medium (Middlebrook 7H9 medium, Difco) supplemented with 10% albumin-dextrose-catalase (ADC; Difco) and 0.05% Tween 80 (Sigma-Aldrich), or on complete 7H11 solid medium (Middlebrook 7H11 agar medium, Difco) supplemented with 10% oleic acid–ADC (OADC; Difco). When required, mycobacterial growth media were supplemented with Km (50 μg ml^{-1}), hygromycin (Hm; 50 μg ml^{-1}), Sm (25 μg ml^{-1}), zeocin (Zc; 25 μg ml^{-1}), Ace (0.2% w/v), or Atc (100 or 200 ng ml^{-1}).

Plasmid constructs

Plasmids pMPMK6 (*39*), p29SEN (*40*), pGMCS (*41*), pGMCZ (*42*), pLAM12 (*43*), pETDuet-1, pET15b and pRARE (Novagen), pBAD30 (*44*), and pTA100 (*4*) have been described. Primers used for plasmid construction are described in table S1. All the plasmids constructed in this work have been verified by sequencing. The pMPMK6 derivatives expressing the toxins, namely, pK6-$MenT_1$, pK6-$MenT_2$, pK6-$MenT_3$, and pK6-$MenT_4$, were constructed as follows: *menT*$_1$, *menT*$_2$, *menT*$_3$, and *menT*$_4$ were PCR-amplified from the *M. tuberculosis* H37Rv genome and cloned as Eco RI/Hind III fragments (*menT*$_1$ and *menT*$_2$) and Mfe I/Hind III fragments (*menT*$_3$ and *menT*$_4$) into Eco RI/Hind III–digested pMPMK6.

The p29SEN plasmid derivatives encoding the antitoxins, namely, p29SEN-$MenA_1$, p29SEN-$MenA_2$, p29SEN-$MenA_3$, and p29SEN-$MenA_4$, were constructed as follows: *menA*$_1$, *menA*$_2$, *menA*$_3$, and *menA*$_4$ were PCR-amplified from the *M. tuberculosis* H37Rv genome and cloned either as Eco RI/Hind III fragments (*menA*$_1$, *menA*$_2$, and *menA*$_3$) or as Mfe I/Hind III fragments (*menA*$_4$) into Eco RI/Hind III–digested p29SEN. For p29SEN-Rph, the *rph* gene was PCR-amplified from the *E. coli* DLT1900 genome and cloned as an Eco RI/Hind III fragment into Eco RI/Hind III–digested p29SEN.

To construct pGMC-$MenT_2$, pGMC-$MenT_3$, and pGMC-$MenT_4$, *menT*$_2$, *menT*$_3$, and *menT*$_4$ were PCR-amplified using pK6-$MenT_2$, pK6-$MenT_3$, and pK6-$MenT_4$ templates, respectively, and cloned into pGMCS using In-Fusion HD Cloning Kits (Takara Bio). Plasmid pGMC-$MenT_1$ and pGMC-$MenT_{1\text{-}His}$ were obtained following PCR amplification of *menT*$_1$ and *menT*$_{1\text{-}His}$ using pK6-$MenT_1$ as a template and homologous recombination in linearized pGMCS plasmid by In-Fusion HD Cloning Kits (Takara Bio). For pGMC-*$MenA_4$-$MenT_4$, the *menA*$_4$-*menT*$_4$ operon was PCR-amplified from the H37Rv genome and cloned into linearized pGMCS plasmid by In-Fusion HD Cloning Kits (Takara Bio).

To construct plasmids pLAM-$MenA_2$, pLAM-$MenA_3$, and pLAM-$MenA_4$, *menA*$_2$, *menA*$_3$, and *menA*$_4$ were PCR-amplified using p29SEN-$MenA_2$, p29SEN-$MenA_3$, and p29SEN-$MenA_4$ as templates, respectively. These were cloned as Nde I/Eco RI fragments (*menA*$_2$ and *menA*$_3$) and Nde I/Mfe I fragments (*menA*$_4$) into Nde I/Eco

RI–digested pLAM12. Plasmid p29SEN-MenA$_1$ was used to amplify *menA$_1$* and *menA$_{1\text{-}His}$*, which were then cloned as Nde I/Eco RI fragments into Nde I/Eco RI–digested pLAM12 to produce pLAM-MenA$_1$ and pLAM-MenA$_{1\text{-}His}$, respectively.

The pET vector derivatives used in this work were constructed as follows. To construct plasmid pET-MenT$_{3\text{-}His}$, *menT$_{3\text{-}His}$* (with an added fragment encoding a Ser-Ser-Gly-His$_6$ C-terminal tag) was PCR-amplified from pK6-MenT$_3$ template and cloned as an Nde I/Mfe I fragment into Nde I/Mfe I–digested pETDuet-1. Plasmid pET-MenT$_{3\text{-}His}$ was used as a template to construct pET-MenT$_{3\text{-}His}^{(D80A)}$ and pET-MenT$_{3\text{-}His}^{(K189A)}$ by QuikChange site-directed mutagenesis (Agilent) using appropriate primers. Plasmid pET-MenA$_{3\text{-}His}$, encoding an N-terminal His$_6$-tagged MenA$_3$ antitoxin, was constructed by PCR amplification of *menA$_{3\text{-}His}$* using p29SEN-MenA$_3$ as a template, Nde I/Hind III digestion, and cloning into Nde I/Hind III–digested pET15b plasmid. To construct plasmid pET-MenT$_3$/MenA$_{3\text{-}His}$, *menA$_{3\text{-}His}$* was first PCR-amplified from p29SEN-MenA$_3$ template and cloned as an Nco I/Hind III fragment into Nco I/Hind III–digested pETDuet-1. *menT$_3$* was then PCR-amplified from pK6-MenT$_3$, digested with Nde I/Mfe I, and cloned into Nde I/Mfe I–digested pET-MenA$_{3\text{-}His}$. To construct pET-MenT$_{3\text{-}His}$/MenA$_3$, *menA$_3$* was first PCR-amplified using p29SEN-MenA$_3$ as a template and cloned as an Nco I/Hind III fragment into Nco I/Hind III–digested pET-MenT$_{3\text{-}His}$. To generate pET-MenT$_{1\text{-}His}$ (expressing MenT$_1$ with an N-terminal His$_6$-Ser-Ser-Gly-tag), *menT$_{1\text{-}His}$* was PCR-amplified from pK6-MenT$_1$ and cloned as an Nde I/Mfe I fragment into Nde I/Mfe I–digested pETDuet-1. For pET-MenA$_{1\text{-}His}$ (expressing MenA$_1$ with an N-terminal His$_6$-Ser-Ser-Gly-tag), *menA$_{1\text{-}His}$* was PCR-amplified from p29SEN-MenA$_1$ template and cloned as an Nco I/Bam HI fragment into Nco I/Bam HI–digested pETDuet-1. For pET-MenT$_1$/MenA$_{1\text{-}His}$, *menT$_1$* was PCR-amplified from pK6-MenA$_1$ and cloned as an Nde I/Mfe I fragment into Nde I/Mfe I–digested pET-MenA$_{1\text{-}His}$. For pET-MenT$_{1\text{-}His}$/MenA$_1$, *menA$_1$* was PCR-amplified from p29SEN-MenA$_1$ and cloned as an Nco I/Bam HI fragment into Nco I/Bam HI–digested pET-MenT$_{1}$His.

To generate MenT$_3$ and MenT$_4$ expression constructs for crystallization and biochemistry, overlap PCRs were performed to fuse a sentrin protease (SENP)–cleavable N-terminal His$_6$-SUMO tag, amplified from the pBAT4 derivative (*45*), pSAT1-LIC (this study), to either *menT$_3$* or *menT$_4$*, amplified from H37Rv genomic DNA. The resulting PCR products were cloned as either Kpn I/Hind III fragments into Kpn I/Hind III–digested pBAD30 (*menT$_3$*), producing pTRB517, or as Xma I/Hind III fragments into Xma I/Hind III–digested pBAD30 (*menT$_4$*) to generate pTRB544.

Plasmids pPF656 and pPF657 were constructed by amplifying *menA$_3$* and *menT$_3$* from H37Rv genomic DNA and cloning as Mfe I/Xma I fragments into Eco RI/Xma I–digested pTA100 and pBAD30, respectively. To express His$_6$-SUMO-tagged MenT$_3^{(D80A)}$, site-directed mutagenesis was carried out using pTRB517 as a template. Briefly, nonoverlapping inverse primers were used to amplify *menT$_3^{(D80A)}$*, followed by incubation with a mix of T4 DNA ligase, T4 polynucleotide kinase, and DpnI at 37°C to remove template and circularize amplified DNA. This reaction was then used to transform *E. coli* DH5α, resulting in pTRB593. Similarly, this method was used to generate MenT$_3^{(D80A)}$, MenT$_3^{(K189A)}$, and MenT$_3^{(D211A)}$ for functional testing, using pPF657 as a template, resulting in pTRB591, pTRB562, and pTRB592, respectively.

Plasmid pTRB491 was generated by amplifying *menA$_3$* from H37Rv genomic DNA and cloning into pSAT1-LIC via ligation-independent cloning (LIC). The pSAT1-LIC plasmid features a LIC site that fuses an N-terminal His$_6$-SUMO tag to the target protein. To produce MenT$_3^{(K189A)}$ protein, the mutated gene was amplified from pTRB562 and similarly cloned into pTRB550 via LIC, resulting in pTRB577. The pTRB550 plasmid features a His$_6$-SUMO LIC site, originally amplified from pSAT1-LIC and cloned as an Eco RI/Hind III fragment into Eco RI/Hind III–digested pBAD30.

To produce plasmids for use in *M. tuberculosis*, *menA$_3$*, *menT$_3$*, or both genes were amplified by PCR using PrimeSTAR GXL DNA polymerase, with *M. tuberculosis* H37Rv genomic DNA as template and primer pairs clo-RBS1-MenA$_3$-attB2/clo-MenA$_3$-attB3, clo-RBS1-MenT$_3$-attB2/clo-MenT$_3$-attB3, clo-RBS4-MenT$_3$-attB2/clo-MenT$_3$-attB3, or clo-RBS1-MenA$_3$-attB2/clo-MenT$_3$-attB3, respectively (tables S1 and S2). RBS1 (AGGAAGACAGGCTGCCC) and RBS4 (ACGAAGACAGGCTGCCC), corresponding to a strong or weak Shine-Dalgarno sequence, respectively, were placed upstream from the ATG translation start of MenA$_3$ or the GTG translation start of MenT$_3$. Plasmids pGMCS-TetR-P1-RBS1-MenA$_3$, pGMCS-TetR-P1-RBS1-MenA$_3$-MenT$_3$, pGMCS-TetR-P1-RBS1-MenT$_3$, or pGMCS-TetR-P1-RBS4-MenT$_3$ were constructed by multisite gateway recombination (*18*), using plasmid pDE43-MCS as the destination vector. These plasmids are integrative vectors (insertion at the attL5 mycobacteriophage insertion site in the *glyV* tRNA gene) and express MenA$_3$, MenT$_3$, or MenA$_3$-MenT$_3$ under the control of P1 (P$_{mycl\ tetO}$), a tetracycline-inducible promoter (table S2) (*46*).

Construction of MenT$_3$ D80A, D211A, and K189A substitutions for use in *M. tuberculosis* was performed as follows: Plasmid pGMCS-TetR-P1-RBS4-MenT$_3$ was amplified by PCR with PrimeSTAR GXL DNA polymerase and the oligonucleotides pairs InFus-MenT$_3$D80A-right/InFus-MenT$_3$D80A-left, InFus-MenT$_3$D211A-right/InFus-MenT$_3$D211A-left, or InFus-MenT$_3$K189A-right/InFus-MenT$_3$K189A-left (table S1). The amplified linear fragments were purified on agarose gels and circularized using the In-Fusion HD Cloning Kit (Takara), as recommended by the manufacturer. Plasmids used to transform Stellar recipient cells were verified by sequencing and introduced by electroporation into *M. tuberculosis* Δ(*menA$_3$-menT$_3$*)::*dif6*/pGMCZ (see the next paragraph).

Construction of *M. tuberculosis* mutants

Mutant strains of *M. tuberculosis* H37Rv were constructed by allelic exchange using recombineering (*43*), as previously described (fig. S2) (*47*). Two ~0.5-kb DNA fragments flanking the *menA$_3$-menT$_3$* operon were amplified by PCR using PrimeSTAR GXL DNA polymerase (Takara), *M. tuberculosis* H37Rv genomic DNA, and the primer pairs MenA$_3$Am-For/MenA$_3$Zc-Am-Rev or MenT$_3$Zc-Av-For/MenT$_3$Av-Rev, respectively (table S1). A three-fragment PCR fused these two fragments to a Zc-resistance cassette flanked by two *dif6* variants of the *M. tuberculosis dif* site and the recombination substrate was recovered by agarose gel purifications. The recipient strain for recombineering was a derivative of *M. tuberculosis* H37Rv carrying two plasmids: pJV53H, an Hm-resistant pJV53-derived plasmid expressing recombineering enzymes (*43*), and the integrative plasmid pGMCS-P1-MenA$_3$, constitutively expressing *menA$_3$* (table S2). This strain was grown in complete 7H9 medium supplemented with Hm until mid-log phase and expression of recombineering enzymes was induced by Ace (0.2%) overnight at 37°C. After induction, electrotransformation was performed with 100 ng of the linear DNA fragment for allelic exchange. After a 48-hour incubation at 37°C, mycobacteria were plated onto agar supplemented

with Zc. Zc-resistant clones were restreaked on the same medium, grown in complete 7H9 without antibiotic, and verified to be carrying the expected allele replacement by PCR amplification of chromosomal DNA and subsequent DNA sequencing, using primers $MenA_3$Am-For/$MenT_3$Av-Rev (fig. S1C and table S1). Spontaneous loss of the Zc-resistance cassette by XerCD-dependent recombination and of the pJV53H plasmid was obtained by serial rounds of culture without antibiotics and phenotypic tests for Zc^S and Hm^S. Plasmid pGMCS-P1-$MenA_3$ was then removed by transformation with pGMCZ, a similar integrative vector but carrying resistance to Zc, resulting in the deleted strain *M. tuberculosis* Δ(*$menA_3$-$menT_3$*)::*dif*6/pGMCZ.

E. coli multicopy plasmid library

E. coli MC4100 Δ*dnaKdnaJ*::Km^R Δ*tig*:Cm^R double mutant (*40*) was partially digested with Sau3 AI restriction enzyme and DNA fragments of about 1.5 to 4 kb in size were purified, then ligated into linearized and dephosphorylated Bam HI–digested pMPM2 (ColE1 origin) plasmid (*39*), and used to transform *E. coli* DH10B. About 25,000 independent transformants were pooled to constitute the multicopy library. This library has previously been used as a tool to identify multicopy suppressors of chaperone mutants (*48*).

Bacterial growth assays

In vivo toxicity and antitoxicity assays by cognate or noncognate antitoxins in *E. coli* were performed as follows. *E. coli* DLT1900 were cotransformed with pMPMK6-vector, pK6-$MenT_1$, -$MenT_2$, -$MenT_3$, or -$MenT_4$ (toxins), and p29SEN-vector, p29SEN-$MenA_1$, -$MenA_2$, -$MenA_3$, or -$MenA_4$ (antitoxins). Transformants were re-seeded from overnight cultures and grown at 37°C to mid-log phase in LB supplemented with Km and Ap, and then serially diluted and spotted on LB-agar plates supplemented with Km and Ap, with or without L-ara (0.1%) and/or IPTG (200 μM). Plates were incubated at 37°C overnight and then imaged and counted. $MenT_3$ substitutions were tested for toxicity in *E. coli* DH5α carrying pBAD30-vector, -$MenT_3$ WT (pPF657), -$MenT_3^{(D80A)}$ (pTRB591), -$MenT_3^{(K189A)}$ (pTRB562), or -$MenT_3^{(D211A)}$ (pTRB592). Strains were grown to mid-log phase, then serially diluted, and spotted onto M9M-agar plates supplemented with Ap, with or without L-ara (0.1%). After a 2-day incubation at 37°C, plates were imaged and counted.

In vivo toxicity and rescue assays by cognate or noncognate antitoxins in *M. smegmatis* were performed as follows. Cultures of mc^2 155 strain grown in LB at 37°C were cotransformed with the integrative pGMC-vector, -$MenT_1$, -$MenT_2$, -$MenT_3$, or -$MenT_4$ (toxins), and with pLAM12-vector, pLAM-$MenA_1$, -$MenA_2$, -$MenA_3$, or -$MenA_4$ (antitoxins). Samples were selected on LB-agar plates supplemented with Km and Sm for 3 days at 37°C, in the presence or absence of Atc (100 ng ml^{-1}) and Ace (0.2%) for toxin and antitoxin expression, respectively. A similar procedure was applied for pGMC-*$MenA_4$-$MenT_4$ carrying the *$menA_4$-$menT_4$* operon, with the exception that no cotransformation with pLAM12 derivatives or selection on Km was needed.

Viability staining and flow cytometry

Exponentially growing cultures [OD_{600} (optical density at 600 nm) between 0.05 and 0.2] of *M. smegmatis* strain mc^2 155 containing plasmid pGMCS-TetR-P1-RBS1-$MenT_3$ were divided in two: Half was left in complete 7H9 growth medium with Sm (uninduced cultures), while the other half was additionally treated with Atc (200 ng ml^{-1}) to induce expression from the P1 promoter. For labeling with LIVE/DEAD BacLight (Molecular Probes) dyes, cells were harvested 8 hours after Atc induction. Cells were centrifuged, resuspended in phosphate-buffered saline buffer, and stained as recommended by the manufacturer. Labeled cells were analyzed by fluorescence-activated cell sorting using a BD LSRFortessa X20 flow cytometer. Flow cytometry data analysis was performed using FlowJo software.

Toxicity assays in *M. tuberculosis*

M. tuberculosis strains H37Rv or H37Rv Δ(*$menA_3$-$menT_3$*)::*dif*6/pGMCZ were transformed by electroporation with 100 ng of plasmids pGMCS-TetR-P1-RBS1-$MenA_3$, pGMCS-TetR-P1-RBS1-$MenA_3$-$MenT_3$, pGMCS-TetR-P1-RBS1-$MenT_3$, pGMCS-TetR-P1-RBS4-$MenT_3$, pGMCS-TetR-P1-RBS4-$MenT_3$(D80A), pGMCS-TetR-P1-RBS4-$MenT_3$(K189A), or pGMCS-TetR-P1-RBS4-$MenT_3$(D211A). After 3 days of phenotypic expression in 7H9 ADC Tween at 37°C, the transformation mix was divided into two halves. One half was plated on 7H11 OADC with Sm; the other half was plated on 7H11 OADC Sm supplemented with Atc (200 ng ml^{-1}). Plates were imaged after 20 days of incubation at 37°C.

In vivo coaffinity purification assays

To perform in vivo copurification assays, *E. coli* BL21 Δ*slyD* was transformed with (i) pET-$MenT_{3\text{-His}}$, pET-$MenA_{3\text{-His}}$, pET-$MenT_3$/$MenA_{3\text{-His}}$, or pET-$MenT_{3\text{-His}}$/$MenA_3$, or with (ii) pET-$MenT_{1\text{-His}}$, pET-$MenA_{1\text{-His}}$, pET-$MenT_1$/$MenA_{1\text{-His}}$, or pET-$MenT_{1\text{-His}}$/$MenA_1$, and selected on LB-agar plates supplemented with Ap and glu (20%). Transformants were grown at 37°C to an OD_{600} of approximately 0.4 and then protein expression was induced overnight at 20°C with 1 mM IPTG. Cell lysis and affinity purification of the protein complexes were performed as described below for $MenT_{3\text{-His}}$ purification. Elution fractions were separated on SDS-PAGE and proteins revealed using InstantBlue Protein Stain (Expedeon, catalog no. ISB1L).

Recombinant protein production

To purify $MenT_3$ for biochemistry, BL21 (λDE3) Δ*slyD* transformed with pET-$MenT_{3\text{-His}}$, pET-$MenT_{3\text{-His}}^{(D80A)}$, or pET-$MenT_{3\text{-His}}^{(K189A)}$ was grown to an OD_{600} of approximately 0.4 at 37°C. IPTG (1 mM) was then added, and the culture was incubated overnight at 20°C. Under such conditions, $MenT_3$ expression in *E. coli* was better tolerated and led to a reasonable amount of soluble $MenT_3$ that could be collected for purification. Cultures were centrifuged at 5000*g* for 10 min at 4°C, pellets were resuspended in Lysis buffer [300 mM NaCl, 50 mM tris (pH 7.5), and protease inhibitor tablet (Roche); 20 ml of buffer per 1 liter of cell culture] and incubated for 30 min on ice. Lysis was performed using the One Shot cell disrupter at 1.5 kbar (One Shot model, Constant Systems Ltd.). Lysates were centrifuged for 30 min at 30,000*g* in 4°C, and the resulting supernatants were gently mixed at 4°C for 30 min with Ni–nitrilotriacetic acid agarose beads (Qiagen, catalog no. 30230) preequilibrated with buffer PD [300 mM NaCl and 50 mM tris (pH 7.5)], using a 10-ml poly-prep column (Bio-Rad, catalog no. 7311550). Columns were stabilized for 10 min at 4°C and washed three times with 10 ml of buffer PD plus 25 mM imidazole, and proteins were then eluted with buffer PD containing 250 mM imidazole. Elutions (500 μl) were collected and PD MiniTrap G-25 columns (GE Healthcare, catalog no. 16924748) were used to exchange buffer with buffer PD

supplemented with 10% glycerol. Proteins were concentrated using Vivaspin 6 columns with a 5000-Da cutoff (Sartorius, catalog no. 184501257). Proteins were stored at −80°C until further use.

For additional $MenT_3$ and $MenT_3^{(K189A)}$ expression, either for crystallization or biochemistry, *E. coli* ER2566 pRARE pPF656 was transformed with either pTRB517 or pTRB577, respectively. For $MenT_3^{(D80A)}$ expression, *E. coli* ER2566 pRARE was transformed with pTRB593. $MenT_4$ was expressed in *E. coli* BL21 (λDE3) transformed with pTRB544. $MenA_3$ was expressed in *E. coli* ER2566 transformed with pTRB491. For these expressions, the same procedure was followed: Overnight cultures were re-seeded 1:100 into 2-liter flasks containing 1-liter 2× YT. Cells were grown at 175 rpm in 37°C until an OD_{600} of 0.3 was reached and then at 22°C until OD_{600} 0.5, whereupon expression was induced by the addition of L-ara (0.1%) for toxins and IPTG (1 mM) for antitoxins. Cells were left to grow overnight at 16°C, shaking at 175 rpm.

For selenomethionine incorporation, starter cultures of ER2566 pRARE pPF656 pTRB517 were grown overnight in LB at 37°C with 200 rpm shaking. Cells were pelleted, washed, and resuspended in M9M, and then sub-cultured into 500 ml of M9M in 2-liter baffled flasks to a starting OD_{600} of 0.075. Cells were grown at 37°C with 175 rpm shaking until an OD_{600} of 0.6, whereupon cells were centrifuged at 4200*g* and resuspended in fresh M9M. This sample was divided between separate 2-liter baffled flasks containing new M9M and shaken at 175 rpm for a further 1 hour at 37°C. Once an OD_{600} of 0.7 was reached, 12 ml of nutrient mix [L-lysine hydrate (4 mg ml^{-1}), L-threonine (4 mg ml^{-1}), L-phenylalanine (4 mg ml^{-1}), L-leucine (2 mg ml^{-1}), L-isoleucine (2 mg ml^{-1}), L-valine (2 mg ml^{-1}), and 4 mM $CaCl_2$] was added to each flask to promote feedback inhibition of methionine synthesis, followed by 250× SelenoMethionine Solution (Molecular Dimensions) to a final concentration of 40 μg ml^{-1}, and cells were left to incubate for 1 hour at 20°C. Last, toxin and antitoxin expression were induced by the addition of L-ara (0.1%) and IPTG (1 mM), and samples were left to grow overnight at 175 rpm in 16°C.

All five proteins were purified in the same manner. Bacteria were harvested by centrifugation at 4200*g*, and the pellets were resuspended in buffer A500 [20 mM tris-HCl (pH 7.9), 500 mM NaCl, 5 mM imidazole, and 10% glycerol]. Cells were lysed by sonication at 40 kpsi and then centrifuged (45,000*g*, 4°C). The clarified lysate was next passed over a HisTrap HP column (GE Healthcare), washed for 10 column volumes with A500, followed by 10 column volumes of buffer A100 [20 mM tris-HCl (pH 7.9), 100 mM NaCl, 5 mM imidazole, and 10% glycerol], and then eluted directly onto a HiTrap Q HP column (GE Healthcare) with buffer B100 [20 mM tris-HCl (pH 7.9), 100 mM NaCl, 250 mM imidazole, and 10% glycerol]. The Q HP column was transferred to an Äkta Pure (GE Healthcare), washed with 3 column volumes of A100, and then proteins were eluted using a gradient from 100% A100 to 100% buffer C1000 [20 mM tris-HCl (pH 7.9), 1000 mM NaCl, and 10% glycerol]. Fractions containing the protein peak were analyzed by SDS-PAGE, pooled, and incubated overnight at 4°C with hSENP2 SUMO protease to cleave the His_6-SUMO tag from the target protein. The following day, the samples were passed through a second HisTrap HP column and the flow-through fractions containing untagged target protein were collected. These samples were concentrated and run over a HiPrep 16/60 Sephacryl S-200 size exclusion column (GE Healthcare) in buffer S [50 mM tris-HCl (pH 7.9), 500 mM KCl, and 10% glycerol]. Peak fractions were analyzed by SDS-PAGE, pooled, and concentrated. Optimal fractions were separated and either flash-frozen in liquid N_2 for storage at −80°C or dialyzed overnight at 4°C into buffer X [20 mM tris-HCl (pH 7.9), 150 mM NaCl, and 2.5 mM dithiothreitol (DTT)] for crystallographic studies. Crystallization samples were quantified and stored on ice and then either used immediately or flash-frozen in liquid N_2 for storage at −80°C. Frozen crystallization samples still formed usable crystals 15 months after storage.

Protein crystallization

Native and selenomethionine-derivatized $MenT_3$ were concentrated to 12 mg ml^{-1} and $MenT_4$ was concentrated to 6 mg ml^{-1}, all in buffer X (see above). Initial crystallization screens were performed using a Mosquito Xtal3 robot (TTP Labtech) to set 200:100 nl and 100:100 nl protein:condition sitting drops. After initial screening and optimization, both $MenT_3$ protein samples formed thick, six-sided needles in condition G5 [0.2 M calcium acetate hydrate, 0.1 M tris (pH 8.5), and 25% w/v polyethylene glycol 2000 monomethyl ether] of Clear Strategy II HT-96 (Molecular Dimensions). $MenT_4$ formed thin, six-sided needles in the same condition as $MenT_3$. To harvest, 20 μl of condition reservoir was added to 20 μl of cryo buffer [25 mM tris-HCl (pH 7.9), 187.5 mM NaCl, 3.125 mM DTT, and 80% glycerol] and mixed quickly by vortexing; an equal volume of this mixture was then added to the drop. After addition of cryo buffer, crystals were immediately extracted using a nylon loop and flash-frozen in liquid N_2.

Data collection and structure determination

Diffraction data were collected at Diamond Light Source on beamlines I04 ($MenT_3$ native), I03 ($MenT_3$ selenomethionine-derivatized), and I24 ($MenT_4$ native) (Table 1). Single 360° datasets were collected for native $MenT_3$ and $MenT_4$. Two 360° datasets from $MenT_3$ selenomethionine-derivatized crystals measured at the selenium peak (0.9793 Å) were merged using iSpyB (Diamond Light Source). Additional $MenT_3$ selenomethionine-derivatized datasets were collected at selenium high remote (0.9641 Å) and inflection (0.9795 Å) wavelengths. Diffraction data were processed with XDS (*49*), and then AIMLESS from CCP4 (*50*) was used to corroborate the space groups (Table 1). The crystal structure of $MenT_3$ was solved by MAD by providing the SHELX suite in CCP4 with the native and three anomalous $MenT_3$ datasets. The solved starting model for $MenT_3$ was built in REFMAC within CCP4. The crystal structure of $MenT_4$ was solved ab initio using ARCIMBOLDO (*51*). Both models were then iteratively refined and built using PHENIX (*52*) and COOT (*53*), respectively. The quality of the final model was assessed using COOT and the wwPDB validation server. Structural figures were generated using PyMOL (Schrödinger). Comparison against models within the Protein Data Bank (PDB) was performed using DALI (*25*).

Genetic screen for suppressors of toxicity

The following genetic procedure was developed and applied to select for *E. coli* genes that confer resistance to the $MenT_3$ toxin. *E. coli* strain DLT1900 was first transformed with pK6-$MenT_3$ (Km^R) plasmid and transformants were selected at 37°C on LB-agar plates supplemented with Km and glu (0.2%) to repress toxin expression from the *araBAD* promoter of pK6-$MenT_3$. DLT1900 containing pK6-$MenT_3$ was then grown in LB supplemented with Km and glu, transformed with the pMPMA2-based multicopy library of *E. coli* genes, and plated on selective LB-agar supplemented with Km, Ap, and L-ara (0.1%) to induce toxin expression. Plates were incubated for 24 hours at 37°C. A control aliquot of transformants plated on nonselective plates (no L-ara) indicated that the number of

transformants tested during the selection procedure was approximately 60,000. Note that under such conditions, *E. coli* DLT1900 pK6-MenT$_3$ transformed with pMPMA2 empty vector did not produce any colonies on selective plates. We identified 72 toxin-resistant colonies that grew on selective plates after 24 hours, although they were smaller and translucent, indicating that growth inhibition by the toxin is not fully blocked by the suppressors identified. Of the 72 toxin-resistant colonies identified, only 41 were able to grow in culture. Plasmids were extracted from the 41 cultures, used to re-transform DLT1900 pK6-MenT$_3$, and plated as above, to validate growth rescue in the presence of MenT$_3$. Of 41 clones, 18 suppressors passed the second round of selection and were sequenced using the pMPMA2-For and -Rev primers (table S1).

In vitro transcription/translation assays

Assays were performed as previously described (*54*). Briefly, template DNAs of DHFR (P0ABQ4), WaaF-Strep (P37692), and GatZ-Strep (P0C8J8) were used for in vitro transcription/translation coupled assays (PURExpress, New England Biolabs). These were performed according to the manufacturer's instructions, in the presence or absence of the toxin. Following protein synthesis reactions of 2 hours at 37°C, samples were separated on SDS-PAGE and visualized by InstantBlue staining (DHFR) or Western blots using anti-Strep tag antibodies (WaaF-Strep and GatZ-Strep).

Identification of uncharged tRNAs in vivo and in vitro

Prevention of *E. coli* tRNA^{Trp} aminoacylation by MenT$_3$ was monitored using a combination of two previously published methods (*29*, *55*). *E. coli* BL21 (λDE3) transformed with pETDuet or pET-MenT$_3$ was grown at 37°C to OD_{600} 0.1 in M9M, whereupon expression of MenT$_3$ was induced with 1 mM IPTG until an OD_{600} of about 0.4. The bacterial culture (25 ml) was then kept on ice and centrifuged for 10 min at 5000*g* in 4°C. The pellet was resuspended in 0.5 ml of cold 0.3 M sodium acetate (pH 4.5) and 10 mM EDTA and transferred to a precooled 1.5-ml microcentrifuge tube, and 0.5 ml of phenol (equilibrated with the same buffer) was then added. After gentle pipetting, the sample was transferred into phase-lock tubes with an additional 400 μl of cold chloroform. After 30 seconds shaking, the sample was first incubated on ice for 15 min and then centrifuged for 20 min at 20,000*g* in 4°C. The aqueous phase was then transferred to a new cold 1.5-ml tube. Five hundred microliters of cold isopropanol was added and immediately mixed. RNA was precipitated for 1 hour at −20°C, before the sample was centrifuged for 30 min at 20,000*g* in 4°C (*55*). The supernatant was discarded and 1 ml of cold 75% ethanol was carefully added without disturbing the RNA pellet. After further centrifugation for 10 min at 20,000*g* in 4°C, the supernatant was removed and the pellet was air-dried until no ethanol remained. The pellet was then resuspended by vigorously mixing in 20 μl of cold 10 mM sodium acetate (pH 4.5) and 1 mM EDTA. Samples were stored at −80°C. Samples were separated on a denaturing urea acrylamide gel for 3 hours at 100 V in 4°C, as previously described (*29*). Northern blot and visualization with a radiolabeled DNA probe against tRNA^{Trp} was performed as previously described (*56*). Note that to distinguish the band of aminoacylated tRNA from its deacylated counterpart on the Northern blot, a chemically deacylated aliquot of RNA sample prepared from strain containing the empty vector was subjected to alkaline treatment. In this case, 46 μl of tris-HCl (pH 9.0) was added to a 4-μl aliquot of the RNA sample and incubated for 2 hours at 37°C. Fifteen microliters of 0.3 M sodium acetate at pH 4.5 was added and followed by 125 μl of 96% ethanol. RNA was precipitated at −20°C for 1 hour, resuspended, and separated as described above.

For in vitro tRNA charging, in vitro transcription/translation assays were performed as above, using *gatZ* as DNA template. After a 2-hour reaction at 37°C with or without MenT$_3$ toxin (10 μM), tRNA extraction, separation, and visualization were performed as described for the in vivo samples.

In vitro transcription of tRNAs

Labeled tRNAs were prepared by in vitro transcription of PCR templates containing an integrated T7 RNA polymerase promoter sequence. The template for *E. coli* tRNA^{Trp} was made by PCR amplification of chromosomal DNA from strain MG1655 with the primers CC2556 and CC2557 (CC2591 for tRNA^{Trp} without CCA) (table S1). The oligos for *M. tuberculosis* tRNAs are given in table S1. The T7 RNA polymerase in vitro transcription reactions were performed in 25-μl total volume, with a 5-μl nucleotide mix of 2.5 mM ATP, 2.5 mM CTP, 2.5 mM GTP, and 60 μM UTP and 2 to 4 μl of 10 mCi ml^{-1} of radiolabeled UTP [α-P^{32}]. Template (0.1 to 0.2 μg) was used per reaction with 1.5 μl of rRNasin (40 U ml^{-1}) (Promega), 5 μl of 5× optimized transcription buffer (Promega), 2 μl of T7 RNA polymerase (20 U ml^{-1}), and 2.5 μl of 100 mM DTT. Template DNA was removed by the addition of 2 μl of RQ DNase (1 U ml^{-1}) (Promega). Unincorporated nucleotides were removed by G50 spin columns (GE Healthcare) according to the manufacturer's instructions, in a final volume of 30 μl. For *E. coli* tRNA^{Trp}, the transcript reaction was gel-purified on a denaturing 5% acrylamide gel and eluted in 0.3 M sodium acetate for 4 hours overnight at 4°C. The supernatant was removed, ethanol-precipitated, and resuspended in 20 to 30 μl of nuclease-free H_2O.

Nucleotide transfer assays

MenT$_3$ NTase activity was assayed in 10-μl reaction volumes containing 50 mM tris-HCl (pH 9.5), 10 mM $MgCl_2$, and 2.5 mM rNTPs and incubated for 20 min at 37°C. Fresh, uniformly labeled tRNA (0.5 μl) was used per assay, with different dilutions of the protein (1, 0.1, 0.01, and 0.001 mg ml^{-1}) in 50 mM tris-HCl (pH 7.8), 300 mM NaCl, and 10% glycerol. The 10-μl reactions were mixed directly with 10 μl of RNA loading dye (95% formamide, 1 mM EDTA, 0.025% SDS, xylene cyanol, and bromophenol blue), denatured at 90°C, and applied to 5% polyacrylamide-urea gels. The gel was vacuum-dried at 80°C and exposed to a PhosphorImager screen.

In vitro antitoxicity assays

The effect of MenA$_3$ antitoxin was assayed using in vitro-transcribed tRNA^{Ser2} as a substrate. For the coincubation assay, MenT$_3$ (5 μM) and increasing molar ratios of MenA$_3$ were incubated with tRNA^{Ser2} and 2.5 mM CTP in 10-μl reaction volumes containing 50 mM tris-HCl (pH 9.5) and 10 mM $MgCl_2$ for 20 min at 37°C. For the postincubation assay, the reactions were first incubated for 20 min at 37°C with MenT$_3$ alone in 7-μl reaction volumes, then 3 μl containing different concentrations of MenA$_3$ were added, and the reactions were incubated for a further 20 min at 37°C.

MenT$_3$ tRNA screening

The tRNA screening was performed using 0.5 μl of uniformly labeled *M. tuberculosis* tRNAs, all containing the CCA motif. The activity

was tested in 50 mM tris-HCl (pH 9.5), 10 mM $MgCl_2$, and 2.5 mM rCTP in 10-μl reaction volumes and incubated for 20 min at 37°C. The transcripts were incubated with 1 μl of $MenT_3$ (0.1 mg ml^{-1}), or with nuclease-free water as a control. The reaction was stopped with 10 μl of RNA loading dye (95% formamide, 1 mM EDTA, 0.025% SDS, xylene cyanol, and bromophenol blue), denatured at 90°C, and applied to 5% polyacrylamide-urea gels. The gel was vacuum-dried at 80°C and exposed to a PhosphorImager screen.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/31/eabb6651/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. P. Norton, M. A. Mulvey, Toxin-antitoxin systems are important for niche-specific colonization and stress resistance of uropathogenic *Escherichia coli*. *PLOS Pathog.* **8**, e1002954 (2012).
2. S. Helaine, A. M. Cheverton, K. G. Watson, L. M. Faure, S. A. Matthews, D. W. Holden, Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. *Science* **343**, 204–208 (2014).
3. R. Page, W. Peti, Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.* **12**, 208–214 (2016).
4. P. C. Fineran, T. R. Blower, I. J. Foulds, D. P. Humphreys, K. S. Lilley, G. P. Salmond, The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 894–899 (2009).
5. D. C. Pecota, T. K. Wood, Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *J. Bacteriol.* **178**, 2044–2050 (1996).
6. F. Goormaghtigh, N. Fraikin, M. Putrinš, T. Hallaert, V. Hauryliuk, A. Garcia-Pino, A. Sjödin, S. Kasvandik, K. Udekwu, T. Tenson, N. Kaldalu, L. Van Melderen, Reassessing the role of Type II toxin-antitoxin systems in formation of *Escherichia coli* Type II persister cells. *MBio* **9**, e00640–18 (2018).
7. M. H. Pontes, E. A. Groisman, Slow growth determines nonheritable antibiotic resistance in *Salmonella enterica*. *Sci. Signal.* **12**, eaax3938 (2019).
8. M. LeRoux, P. H. Culviner, Y. J. Liu, M. L. Littlehale, M. T. Laub, Stress induces the transcription of toxin-antitoxin systems but does not activate the toxin. *bioRxiv*, (2020).
9. S. Ronneau, S. Helaine, Clarifying the link between toxin–antitoxin modules and bacterial persistence. *J. Mol. Biol.* **431**, 3462–3471 (2019).
10. N. Fraikin, F. Goormaghtigh, L. Van Melderen, Type II toxin-antitoxin systems: Evolution and revolutions. *J. Bacteriol.* **202**, e00763-19 (2020).
11. S. Szekeres, M. Dauti, C. Wilde, D. Mazel, D. A. Rowe-Magnus, Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol. Microbiol.* **63**, 1588–1605 (2007).
12. S. Song, T. K. Wood, Toxin/antitoxin system paradigms: Toxins bound to antitoxins are not likely activated by preferential antitoxin degradation. *Adv. Biosyst.* **4**, 1900290 (2020).
13. A. M. Hall, B. Gollan, S. Helaine, Toxin–antitoxin systems: Reversible toxicity. *Curr. Opin. Microbiol.* **36**, 102–110 (2017).
14. I. Keren, S. Minami, E. Rubin, K. Lewis, Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *MBio* **2**, e00100–11 (2011).
15. A. Sala, P. Bordes, P. Genevaux, Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins* **6**, 1002–1020 (2014).
16. H. Akarsu, P. Bordes, M. Mansour, D.-J. Bigot, P. Genevaux, L. Falquet, TASmania: A bacterial Toxin-Antitoxin Systems database. *PLoS Comput. Biol.* **15**, e1006946 (2019).
17. H. R. Ramage, L. E. Connolly, J. S. Cox, Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: Implications for pathogenesis, stress responses, and evolution. *PLOS Genet.* **5**, e1000767 (2009).
18. D. M. Freire, C. Gutierrez, A. Garza-Garcia, A. D. Grabowska, A. J. Sala, K. Ariyachaokun, T. Panikova, K. S. H. Beckham, A. Colom, V. Pogenberg, M. Cianci, A. Tuukkanen, Y.-M. Boudehen, A. Peixoto, L. Botella, D. I. Svergun, D. Schnappinger, T. R. Schneider, P. Genevaux, L. P. S. de Carvalho, M. Wilmanns, A. H. A. Parret, O. Neyrolles, An NAD+ phosphorylase toxin triggers *Mycobacterium tuberculosis* cell death. *Mol. Cell* **73**, 1282–1291.e8 (2019).
19. P. Garvey, G. F. Fitzgerald, C. Hill, Cloning and DNA sequence analysis of two abortive infection phage resistance determinants from the lactococcal plasmid pNP40. *Appl. Environ. Microbiol.* **61**, 4321–4328 (1995).
20. R. L. Dy, R. Przybilski, K. Semeijn, G. P. C. Salmond, P. C. Fineran, A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. *Nucleic Acids Res.* **42**, 4590–4605 (2014).
21. F. Li, Y. Xiong, J. Wang, H. D. Cho, K. Tomita, A. M. Weiner, T. A. Steitz, Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell* **111**, 815–824 (2002).
22. H. G. Hampton, S. A. Jackson, R. D. Fagerlund, A. I. M. Vogel, R. L. Dy, T. R. Blower, P. C. Fineran, AbiEi binds cooperatively to the Type IV abiE toxin–antitoxin operator via a positively-charged surface and causes DNA bending and negative autoregulation. *J. Mol. Biol.* **430**, 1141–1156 (2018).
23. C. M. Sassetti, D. H. Boyd, E. J. Rubin, Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 (2003).
24. M. A. Dejesus, E. R. Gerrick, W. Xu, S. W. Park, J. E. Long, C. C. Boutte, E. J. Rubin, D. Schnappinger, S. Ehrt, S. M. Fortune, C. M. Sassetti, T. R. Ioerger, Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *MBio* **8**, e02133–16 (2017).
25. L. Holm, C. Sander, Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
26. Y. Zhao, X. Ye, Y. Su, L. Sun, F. She, Y. Wu, Crystal structure confirmation of JHP933 as a nucleotidyltransferase superfamily protein from *Helicobacter pylori* strain J99. *PLOS ONE* **9**, e104609 (2014).
27. S. Bost, F. Silva, D. Belin, Transcriptional activation of *ydeA*, which encodes a member of the major facilitator superfamily, interferes with arabinose accumulation and induction of the *Escherichia coli* arabinose P_{BAD} promoter. *J. Bacteriol.* **181**, 2185–2191 (1999).
28. D. H. Bechhofer, M. P. Deutscher, Bacterial ribonucleases and their roles in RNA metabolism. *Crit. Rev. Biochem. Mol. Biol.* **54**, 242–300 (2019).
29. B. D. Janssen, E. J. Diner, C. S. Hayes, in *Bacterial Regulatory RNA* (Humana Press, 2012; www.ncbi.nlm.nih.gov/pubmed/22736012), vol. 905, pp. 291–309.
30. P. P. Chan, T. M. Lowe, GtRNAdb 2.0: An expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–D189 (2016).
31. L. Samhita, V. Nanjundiah, U. Varshney, How many initiator tRNA genes does *Escherichia coli* need? *J. Bacteriol.* **196**, 2607–2615 (2014).
32. B. Lenhard, O. Orellana, M. Ibba, I. Weygand-Durasević, tRNA recognition and evolution of determinants in seryl-tRNA synthesis. *Nucleic Acids Res.* **27**, 721–729 (1999).
33. G. Jankevicius, A. Ariza, M. Ahel, I. Ahel, The toxin-antitoxin system DarTG catalyzes reversible ADP-ribosylation of DNA. *Mol. Cell* **64**, 1109–1116 (2016).
34. K. S. Winther, D. E. Brodersen, A. K. Brown, K. Gerdes, VapC20 of *Mycobacterium tuberculosis* cleaves the Sarcin–Ricin loop of 23S rRNA. *Nat. Commun.* **4**, 2796 (2013).
35. A. M. Cheverton, B. Gollan, M. Przydacz, C. T. Wong, A. Mylona, S. A. Hare, S. Helaine, A *Salmonella* toxin promotes persister formation through acetylation of tRNA. *Mol. Cell* **63**, 86–96 (2016).
36. D. Jurėnas, S. Chatterjee, A. Konijnenberg, F. Sobott, L. Droogmans, A. Garcia-Pino, L. Van Melderen, AtaT blocks translation initiation by N-acetylation of the initiator $tRNA^{fMet}$. *Nat. Chem. Biol.* **13**, 640–646 (2017).
37. J.-Y. Bouet, M. Bouvier, D. Lane, Concerted action of plasmid maintenance functions: Partition complexes create a requirement for dimer resolution. *Mol. Microbiol.* **62**, 1447–1459 (2006).
38. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640–6645 (2000).
39. M. P. Mayer, A new set of useful cloning and expression vectors derived from pBlueScript. *Gene* **163**, 41–46 (1995).
40. P. Genevaux, F. Keppel, F. Schwager, P. S. Langendijk-Genevaux, F. U. Hartl, C. Georgopoulos, *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**, 195–200 (2004).
41. A. Blumenthal, C. Trujillo, S. Ehrt, D. Schnappinger, Simultaneous analysis of multiple *Mycobacterium tuberculosis* knockdown mutants *in vitro* and *in vivo*. *PLOS ONE* **5**, e15667 (2010).
42. L. Botella, J. Vaubourgeix, J. Livny, D. Schnappinger, Depleting *Mycobacterium tuberculosis* of the transcription termination factor Rho causes pervasive transcription and rapid death. *Nat. Commun.* **8**, 14731 (2017).
43. J. C. van Kessel, G. F. Hatfull, Recombineering in *Mycobacterium tuberculosis*. *Nat. Methods* **4**, 147–152 (2007).
44. L. M. Guzman, D. Belin, M. J. Carson, J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
45. J. Peränen, M. Rikkonen, M. Hyvönen, L. Kääriäinen, T7 vectors with a modified T7*lac* promoter for expression of proteins in *Escherichia coli*. *Anal. Biochem.* **236**, 371–373 (1996).
46. S. Ehrt, X. V. Guo, C. M. Hickey, M. Ryou, M. Monteleone, L. W. Riley, D. Schnappinger, Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res.* **33**, e21 (2005).
47. Y.-M. Boudehen, M. Wallat, P. Rousseau, O. Neyrolles, C. Gutierrez, An improved Xer-cise technology for the generation of multiple unmarked mutants in mycobacteria. *Biotechniques* **68**, 106–110 (2020).

48. F. Anglès, M.-P. Castanié-Cornet, N. Slama, M. Dinclaux, A.-M. Cirinesi, J.-C. Portais, F. Létisse, P. Genevaux, Multilevel interaction of the DnaK/DnaJ(HSP70/HSP40) stress-responsive chaperone machine with the central metabolism. *Sci. Rep.* **7**, 41341 (2017).
49. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
50. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
51. D. D. Rodríguez, C. Grosse, S. Himmel, C. González, I. M. de Ilarduya, S. Becker, G. M. Sheldrick, I. Usón, Crystallographic *ab initio* protein structure solution below atomic resolution. *Nat. Methods* **6**, 651–653 (2009).
52. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
53. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
54. P. Bordes, A. J. Sala, S. Ayala, P. Texier, N. Slama, A. M. Cirinesi, V. Guillet, L. Mourey, P. Genevaux, Chaperone addiction of toxin-antitoxin systems. *Nat. Commun.* **7**, 13339 (2016).
55. T. S. Stenum, M. A. Sørensen, S. Lo Svenningsen, Quantification of the abundance and charging levels of transfer RNAs in *Escherichia coli*. *J. Vis. Exp.* **22**, 56212 (2017).
56. V. Khemici, J. Prados, P. Linder, P. Redder, Decay-initiating endoribonucleolytic cleavage by RNase Y is kept under tight control via sequence preference and sub-cellular localisation. *PLOS Genet.* **11**, e1005577 (2015).

Acknowledgments: We thank D.-J. Bigot for plasmid constructs, and P. Bordes, M.-P. Castanié-Cornet, L. Falquet, L. Poljak, L. Hadjeras, and H. Akarsu for valuable advice. We also thank K. Semeijn and R. Dy for initial plasmid construction and testing, and E. Naser (Genotoul TRI-IPBS imaging facility) for help with flow cytometry analysis. **Funding:** This work was supported by a scholarship from the China Scholarship Council (CSC) as part of a joint international PhD program with Toulouse University Paul Sabatier (Y.C.); Springboard Award (SBF002\1104) from the Academy of Medical Sciences (B.U. and T.R.B.); University of Otago Research Grant (P.C.F.), the School of Biomedical Sciences Bequest Fund, and University of Otago (P.C.F.); CNRS (UPR 9073), Université Paris VII-Denis Diderot, the Agence Nationale de la Recherche (ARNr-QC), and the Labex (Dynamo) program (A.T. and C.C.); European Commission (contracts NEWTBVAC n°241745 and TBVAC2020 n°643381), Centre National de la Recherche Scientifique, Université Paul Sabatier, Agence Nationale de la Recherche (ANR-13-BSV8-0010-01), and Fondation pour la Recherche Médicale (DEQ20160334902) (C.G. and O.N.); and grant SNF CRSII3_160703 (P.G.).
Author contributions: Conceptualization, all authors. Investigation, Y.C., B.U., C.G., A.T., and M. M. Writing, all authors. Funding acquisition, P.C.F., C.C., O.N., P.G., and T.R.B. Supervision, C.C., O.N., P.G., and T.R.B. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The crystal structures of $MenT_3$ and $MenT_4$ have been deposited in the Protein Data Bank under accession numbers 6Y5U and 6Y56, respectively. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 10 March 2020
Accepted 12 June 2020
Published 29 July 2020
10.1126/sciadv.abb6651

Citation: Y. Cai, B. Usher, C. Gutierrez, A. Tolcan, M. Mansour, P. C. Fineran, C. Condon, O. Neyrolles, P. Genevaux, T. R. Blower, A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems. *Sci. Adv.* **6**, eabb6651 (2020).

ScienceAdvances

A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems

Yiming CaiBen UsherClaude GutierrezAnastasia TolcanMoise MansourPeter C. FineranCiarán CondonOlivier NeyrollesPierre GenevauxTim R. Blower

Sci. Adv., 6 (31), eabb6651.

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/31/eabb6651/DC1

Supplementary Materials for

A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems

Yiming Cai, Ben Usher, Claude Gutierrez, Anastasia Tolcan, Moise Mansour, Peter C. Fineran, Ciarán Condon, Olivier Neyrolles, Pierre Genevaux*, Tim R. Blower*

*Corresponding author. Email: genevaux@ibcg.biotoul.fr (P.G.), timothy.blower@durham.ac.uk (T.R.B.)

Published 29 July 2020, *Sci. Adv.* **6**, eabb6651 (2020)
DOI: 10.1126/sciadv.abb6651

This PDF file includes:

SUPPLEMENTARY FIGURES

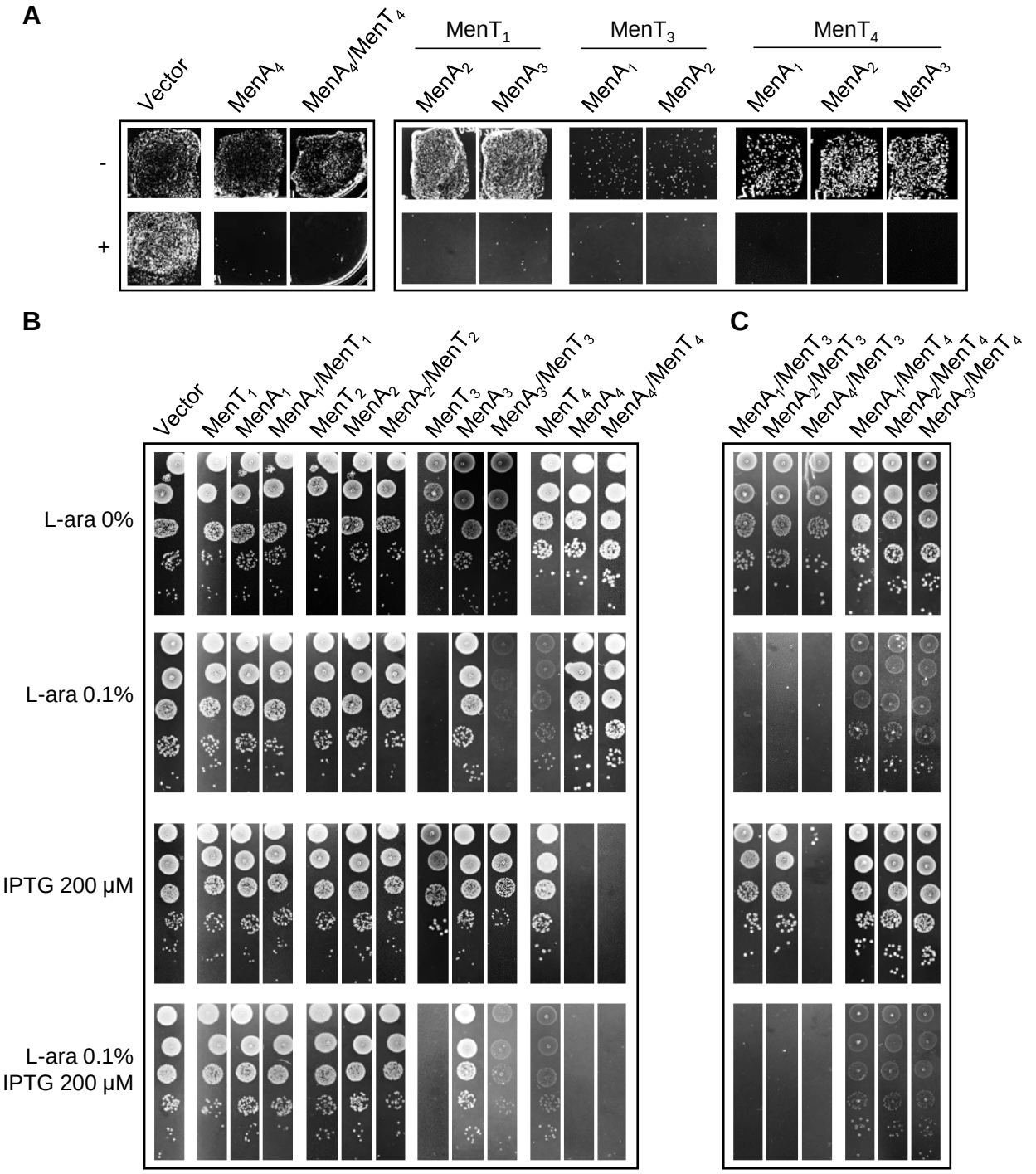


Fig. S1. *M. tuberculosis* TA systems tested for activity and cross-talk in *M. smegmatis* and *E. coli*. (**A**) Toxicity of $MenA_4$ antitoxin and cross-talk between non-cognate toxins and antitoxins. Co-transformants of *M. smegmatis* mc^2 155 containing pGMC-vector, -$MenT_1$, -$MenT_3$ or -$MenT_4$ (toxins), and pLAM-vector, -$MenA_1$, -$MenA_2$, -$MenA_3$, or -$MenA_4$ (antitoxins), were plated on LB-agar in the presence or absence of Atc (100 $ng.ml^{-1}$) and Ace (0.2%) inducers for toxin and antitoxin expression, respectively. Plates were incubated for 3 days at 37 °C; "-" and "+" represent absence or presence of inducer, respectively. (**B**) *E. coli* DLT1900 was co-transformed with the pK6-vector expressing $MenT_1$, $MenT_2$, $MenT_3$ or $MenT_4$ (toxins, L-ara-inducible), in combination with either p29SEN-vector only or p29SEN-expressing the cognate antitoxins $MenA_1$, $MenA_2$, $MenA_3$ or $MenA_4$ (IPTG-inducible). Samples were serial diluted and spotted on agar plates containing L-ara and/or IPTG for toxin and antitoxin expression, respectively. Plates were incubated at 37 °C. (**C**) Cross-talk between the $MenT_3$ and $MenT_4$ toxins that demonstrated toxicity in (B) and the three non-cognate antitoxins, performed as per (B).

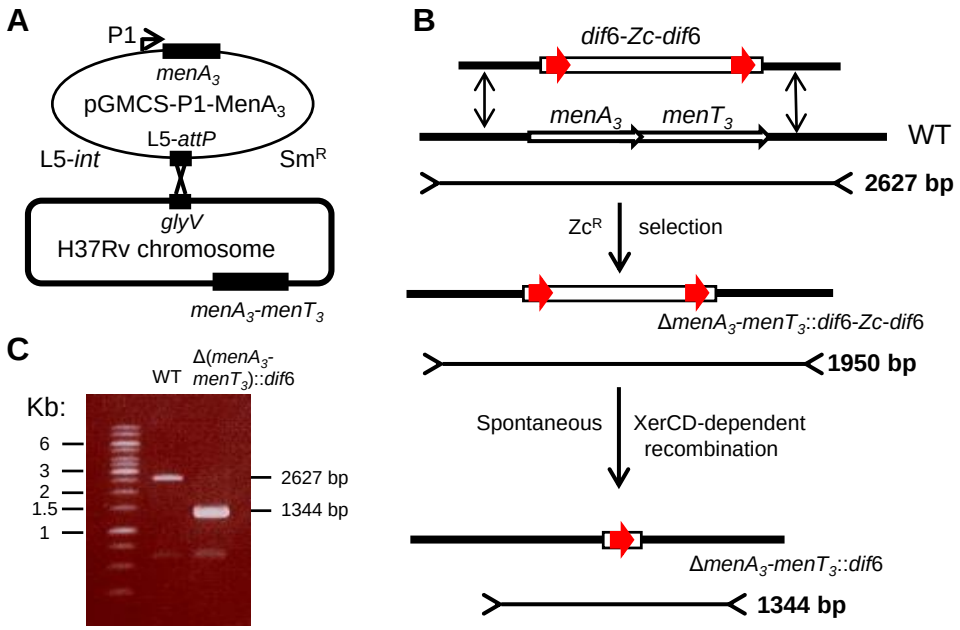


Fig. S2. Construction of a deletion of the *menA$_3$-menT$_3$* operon in *M. tuberculosis* H37Rv. (**A**) A merodiploid strain constitutively expressing an ectopic copy of *menA$_3$* was constructed by integration of plasmid pGMCS-P1-MenA$_3$ (Sm^R). (**B**) Deletion of the operon was obtained in the merodiploid containing plasmid pJV53H by recombineering with a DNA fragment harboring a Zc-resistance cassette flanked by ~500 bp fragments of upstream and downstream regions. Red arrows indicate variants of the *M. tuberculosis dif* site allowing excision of the Zc-resistance gene by spontaneous XerCD-dependent recombination. Replacement of pGMCS-P1-MenA$_3$ by an empty plasmid conferring Zc resistance yielded the H37Rv Δ(*menA$_3$-menT$_3$)::dif6* strain (H37Rv$^{\Delta TA}$). (**C**) Deleted clones were verified by PCR with appropriate oligonucleotides. As indicated in (B), the flanking oligonucleotides amplify a 2627 bp DNA fragment from the WT chromosome and a 1344 bp fragment from that of the deleted strain. These amplicons were verified by DNA sequencing.

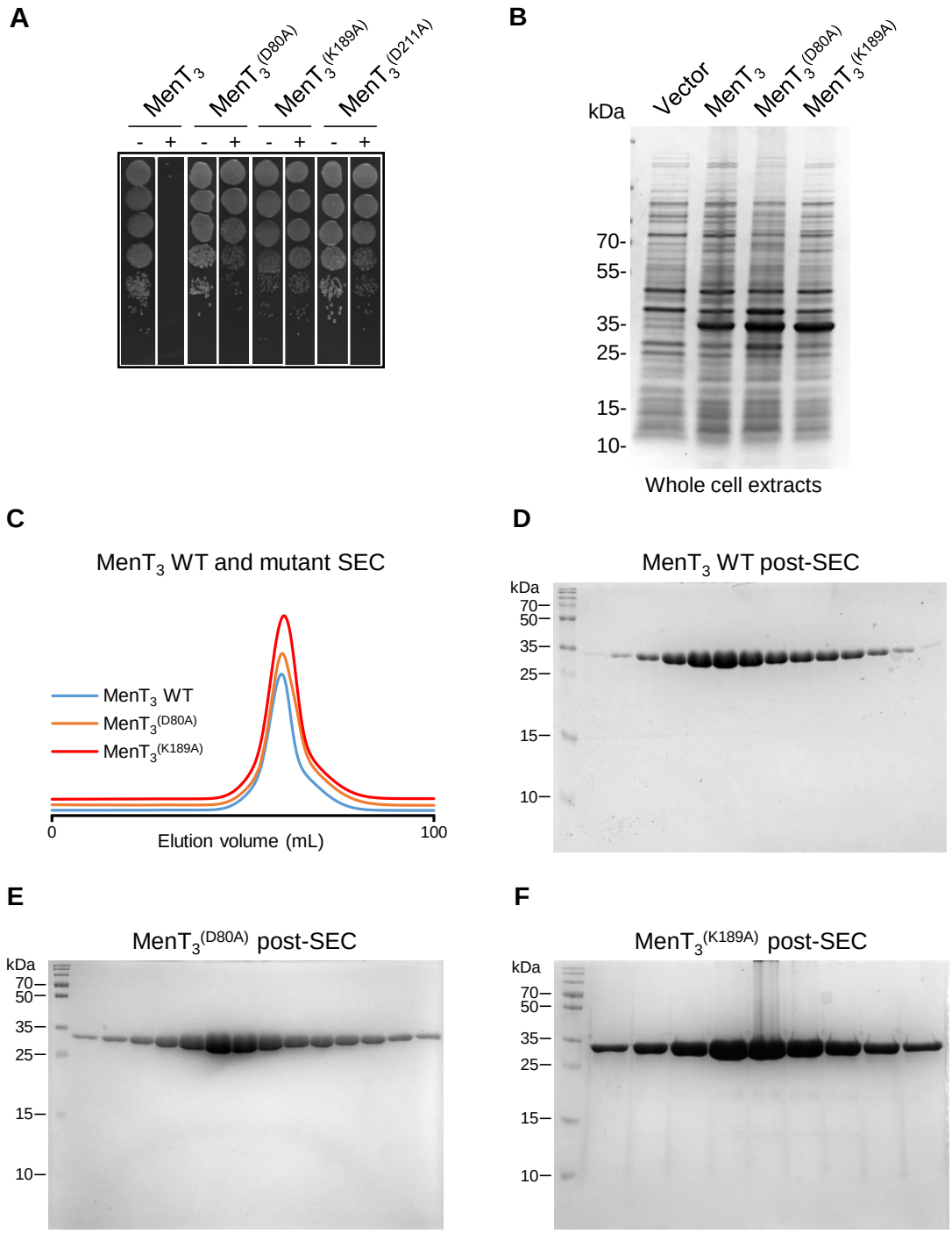
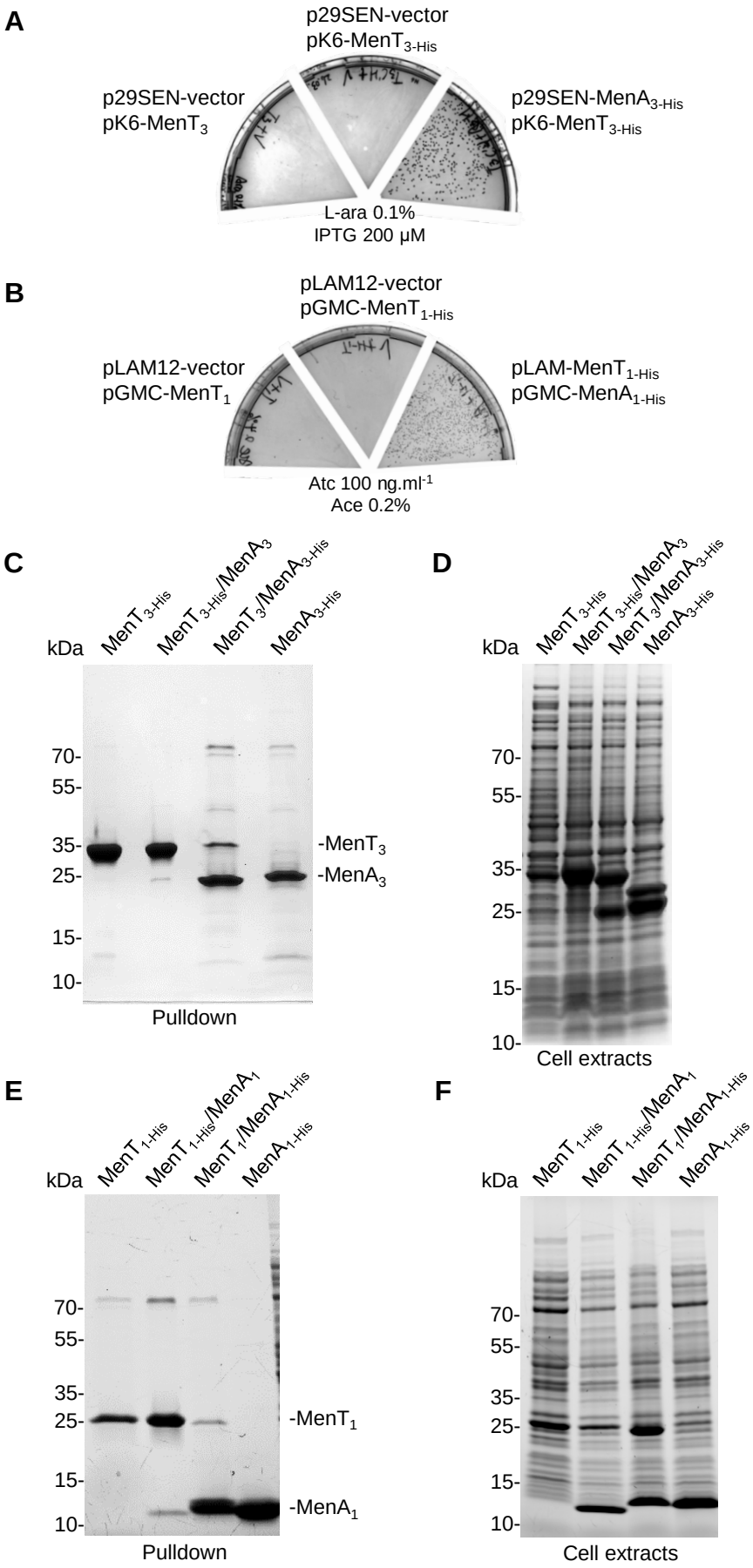


Fig. S3. D80A, K189A and D211A substitutions abolish toxicity of $MenT_3$. (**A**) Toxicity assays of *E. coli* DH5α strains expressing L-ara-inducible $MenT_3$ WT (pPF657) or substituted toxins $MenT_3^{(D80A)}$ (pTRB591), $MenT_3^{(K189A)}$ (pTRB562), and $MenT_3^{(D211A)}$ (pTRB592). Samples were plated on M9M-agar, supplemented with Ap and either glu (0.2%) or L-ara (0.1%), and grown for approximately 48 hrs at 37 °C. "+" and "-" denote expression state. Data are representative of triplicate experiments. (**B**) Coomassie-stained SDS-PAGE of whole cell extracts of *E. coli* BL21 strain expressing $MenT_3$, $MenT_3^{(D80A)}$ or $MenT_3^{(K189A)}$ following a 16 hr incubation at 20°C in the presence of 1mM IPTG. (**C**) Chromatogram showing the comparable elution volumes of $MenT_3$ WT, $MenT_3^{(D80A)}$, and $MenT_3^{(K189A)}$ during size-exclusion chromatography (SEC). All proteins were subjected to the same column, buffers, and conditions. (**D** to **F**) Fractions corresponding to SEC chromatogram peaks were analyzed by SDS-PAGE to verify protein purification. Staining was performed with InstantBlue.

Fig. S4. *In vivo* co-purification of TA complexes. (**A**) To test His-tagged $MenT_3$ and $MenA_3$ for activity, *E. coli* DLT1900 was co-transformed with pK6-vector or pK6-$MenT_{3\text{-}His}$, and p29SEN-vector or p29SEN-$MenA_{3\text{-}His}$. Transformants were plated on LB-agar supplemented with Km and Ap, and in the presence or absence of L-ara (0.1%) and IPTG (200 µM) for toxin and antitoxin expression, respectively. Plates were incubated overnight at 37 °C. (**B**) For His-tagged $MenT_1$ and $MenA_1$, *M. smegmatis* mc^2 155 was co-transformed with pGMC-vector or pGMC-$MenT_{1\text{-}His}$, and pLAM12 or pLAM-$MenA_{1\text{-}His}$. Transformants were plated on LB-agar supplemented with Km and Ap, and in the presence or absence of Atc (100 $ng.ml^{-1}$) and Ace (0.2%) for toxin and antitoxin expression, respectively. Plates were incubated for 3 days at 37 °C. (**C** to **F**) *In vivo* interaction between the toxins and their cognate antitoxins was investigated using *in vivo* co-purification assays in *E. coli*. Cultures of *E. coli* BL21 Δ*slyD* transformed with pET-$MenT_{3\text{-}His}$, pET-$MenA_{3\text{-}His}$, pET-$MenT_3$/$MenA_{3\text{-}His}$, pET-$MenT_{3\text{-}His}$/$MenA_3$, pET-$MenT_{1\text{-}His}$, pET-$MenA_{1\text{-}His}$, pET-$MenT_1$/$MenA_{1\text{-}His}$, or pET-$MenT_{1\text{-}His}$/$MenA_1$, were grown to mid-log phase and protein expression was induced overnight at 20 °C with 1 mM IPTG. Following cell lysis, soluble fractions were applied to a Ni-NTA column, washed, and protein or protein complexes were eluted with 250 mM imidazole. Elution fractions (C and E) and cell extracts (D and F) were separated by SDS-PAGE and stained with InstantBlue.



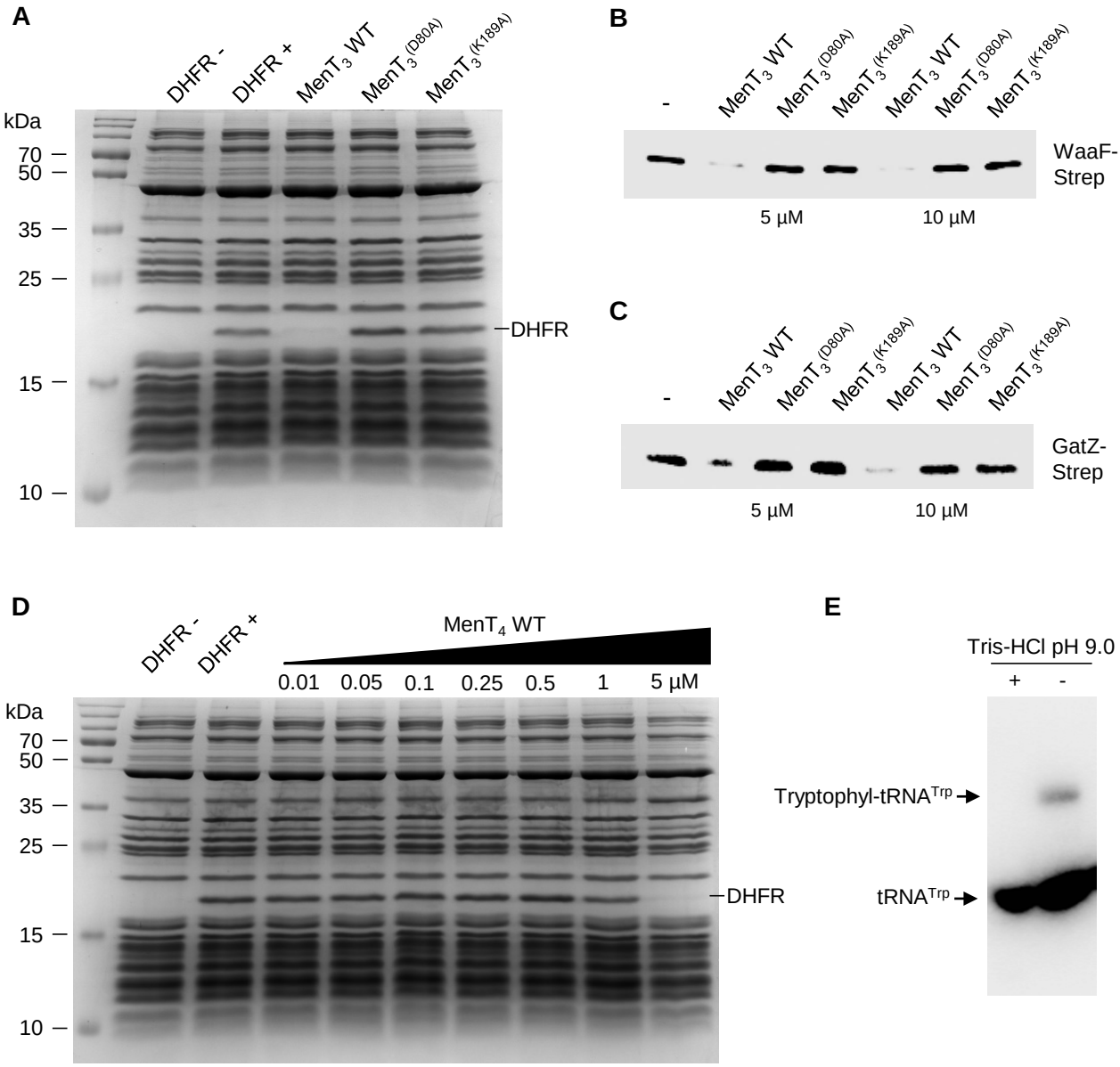


Fig. S5. $MenT_3$ and $MenT_4$ inhibit protein synthesis. (**A**) *In vitro* transcription/translation assays assessing levels of DHFR control protein produced in the absence or presence of 1 µM $MenT_3$ WT or $MenT_3^{(D80A)}$ and $MenT_3^{(K189A)}$ substitutions. Samples were separated by SDS-PAGE. (**B**) and (**C**) WaaF-Strep and GatZ-Strep C-terminal fusion proteins were expressed using an *in vitro* transcription/translation system with or without purified $MenT_3$ WT, $MenT_3^{(D80A)}$ and $MenT_3^{(K189A)}$, at a concentration of 5 or 10 µM. After translation, reactions were separated by SDS-PAGE and target proteins were detected by western blot using anti-Strep antibodies. (**D**) *In vitro* transcription/translation assays assessing levels of DHFR control protein produced in the absence or presence of increasing concentrations of $MenT_4$ WT. Samples were separated by SDS-PAGE. All staining was performed with InstantBlue. (**E**) Base hydrolysis of aminoacylated $tRNA^{Trp}$ extracted *in vivo* from *E. coli* transformed with empty plasmid control. tRNAs were extracted, chemically deacylated following treatment with Tris-HCl pH 9.0, separated, and visualized by Northern blot using specific radiolabeled probes against $tRNA^{Trp}$.

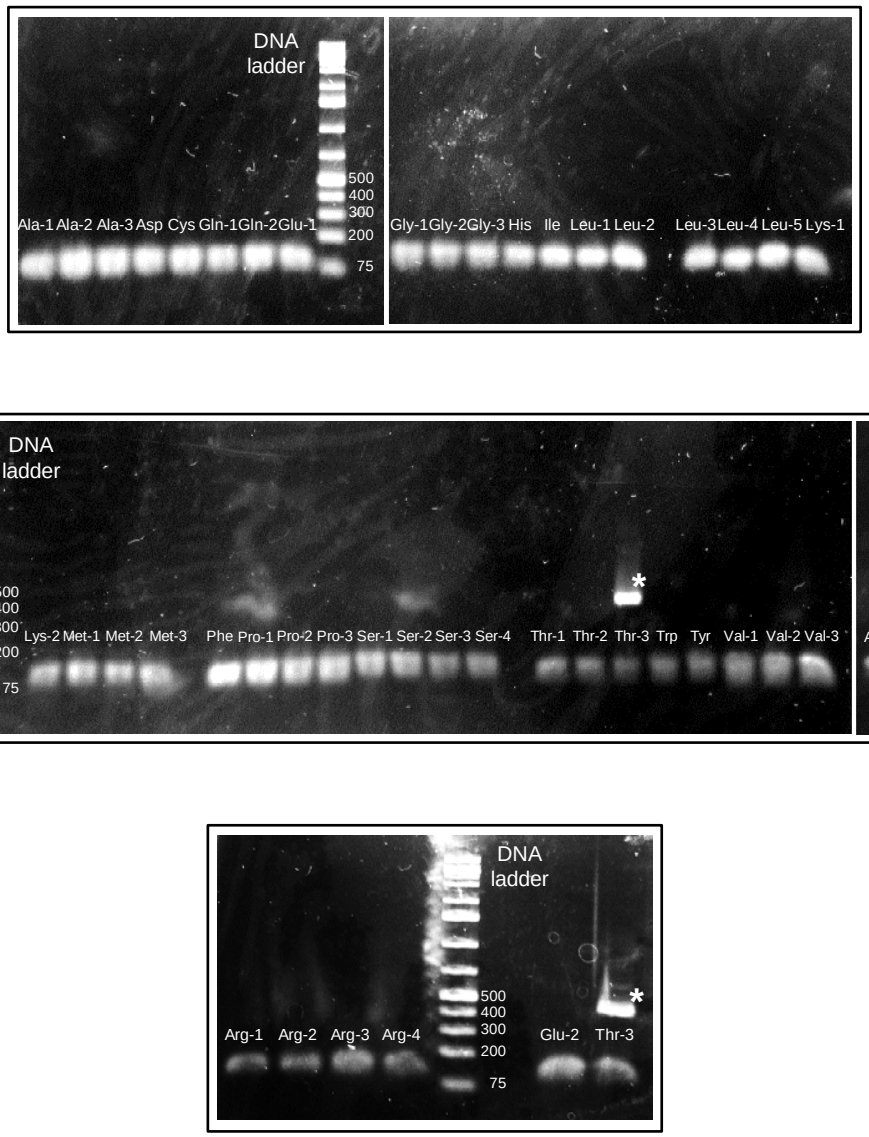


Fig. S6. PCR amplification of all *M. tuberculosis* tRNAs. The 45 *M. tuberculosis* tRNA genes were amplified by PCR from *M. tuberculosis* H37Rv genomic DNA (see **table S1** for primers) and separated on a 1.5 % agarose gel. Note that $tRNA^{Thr3}$ amplification reproducibly showed an upper band after PCR amplification (*), which appeared even after the smaller band of the correct size was extracted, purified and used as a template for a second round of PCR amplification.

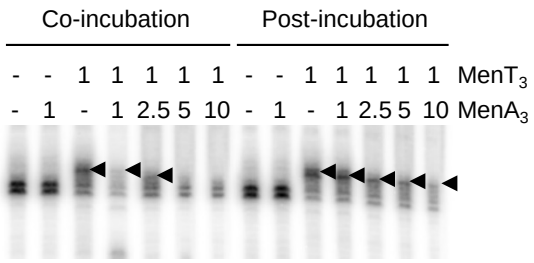


Fig. S7. Co-incubation of $MenA_3$ with $MenT_3$ blocks $MenT_3$ activity *in vitro*, but cannot reverse $MenT_3$ NTase activity by removing nucleotides. Radiolabeled *M. tuberculosis* $tRNA^{Ser2}$ was incubated with 5 μM $MenT_3$ toxin for 20 min at 37 °C, either in the presence of increasing molar ratios of $MenA_3$ antitoxin (co-incubation), or $MenA_3$ was added afterwards (post-incubation) and the reaction continued for a further 20 min at 37 °C. Extended products are indicated with arrowheads.

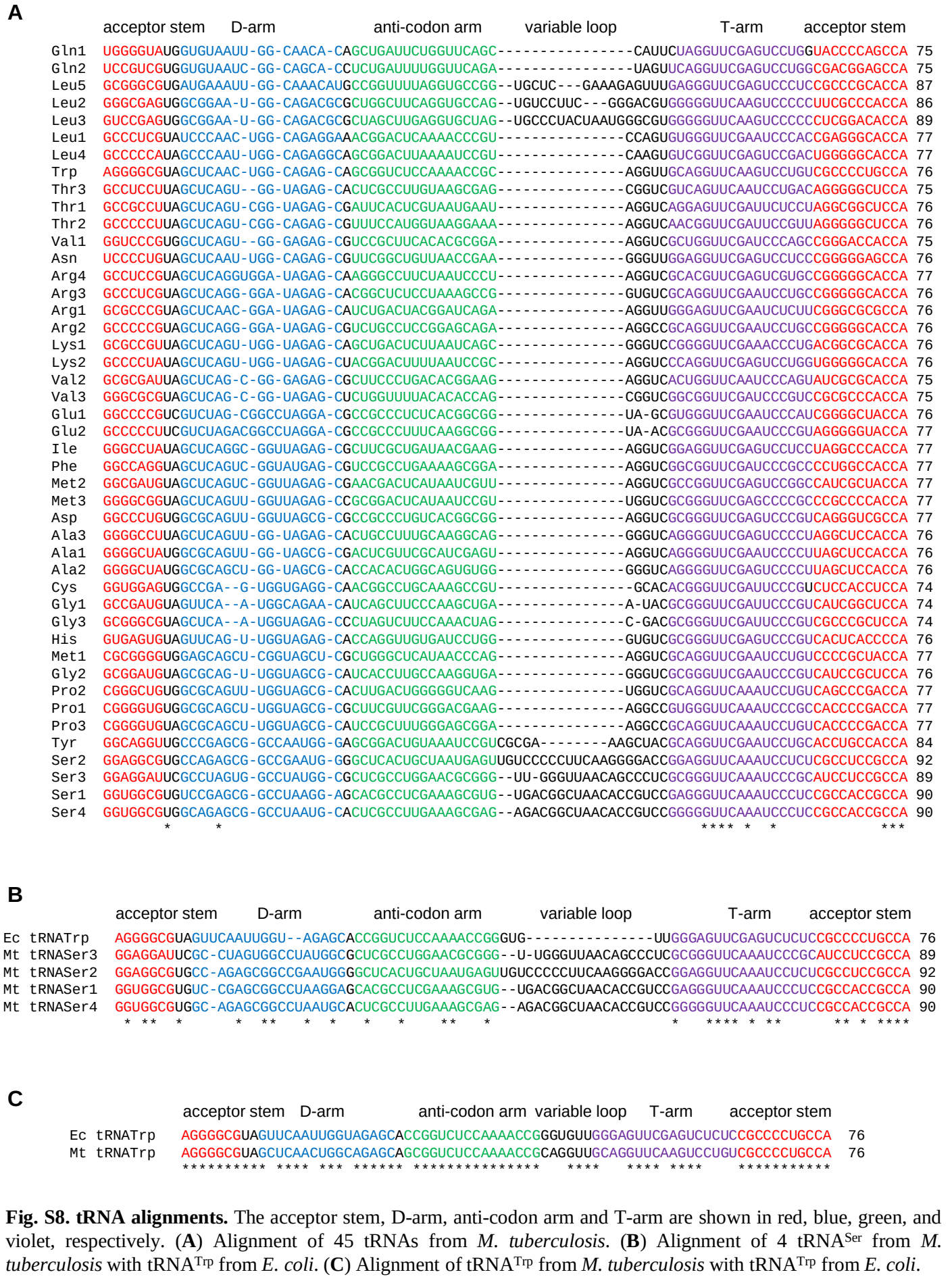


Fig. S8. tRNA alignments. The acceptor stem, D-arm, anti-codon arm and T-arm are shown in red, blue, green, and violet, respectively. (**A**) Alignment of 45 tRNAs from *M. tuberculosis*. (**B**) Alignment of 4 $tRNA^{Ser}$ from *M. tuberculosis* with $tRNA^{Trp}$ from *E. coli*. (**C**) Alignment of $tRNA^{Trp}$ from *M. tuberculosis* with $tRNA^{Trp}$ from *E. coli*.

SUPPLEMENTARY TABLES

Table S1. Oligonucleotides used in this study

Primer name	Sequence (5′-3′)
Construction of *M. tuberculosis* chromosomal mutant	
$MenA_3$Am-For	CCCTGCGAGATGTACTTC
$MenA_3$Zc-Am-Rev	CAGTCGATCCACGTGGAGGATCCCAGATAGCCATGTG
$MenT_3$Zc-Av-For	CCACTGAGCGTCAGACCCACGTGCTCACCTTGAACTCGCCAGGAC
$MenT_3$Av-Rev	GTCACATCCCTTCGACATCTC
Gateway cloning	
clo-RBS1-$MenA_3$-attB2	GGGGACAGCTTTCTTGTACAAAGTGGAGGAAGACAGGCTGCCCATGGCTATCTGGGATCGCCTCG
clo-$MenA_3$-attB3	GGGGACAACTTTGTATAATAAAGTTGTCACGCCGATGCTCGCTTCGGCCGAGCGCG
clo-RBS1-$MenT_3$-attB2	GGGGACAGCTTTCTTGTACAAAGTGGAGGAAGACAGGCTGCCCGTGACCAAGCCCTATTCGTCG
clo-RBS4-$MenT_3$-attB2	GGGGACAGCTTTCTTGTACAAAGTGGACGAAGACAGGCTGCCCGTGACCAAGCCCTATTCGTCG
clo-$MenT_3$-attB3	GGGGACAACTTTGTATAATAAAGTTGTCATCTTTTCGTCGCCCGATCAATCCGCGCAACG
Directed mutagenesis of $MenT_3$ on pGMCS-TetR-P1-$MenT_3$	
InFus-$MenT_3$D80A-right	CGACACGGTCGCACGTCGC
InFus-$MenT_3$D80A-left	CGTGCGACCGTGTCGAAGGCTTTGGAGGTCCGCGAATC
InFus-$MenT_3$D211A-right	GCAGCTTCTTGAAGGACTG
InFus-$MenT_3$D211A-left	CCTTCAAGAAGCTGCAAGGCCACCAGGTCGTGAGCGCGG
InFus-$MenT_3$K189A-right	GCACGCAGTAACTGCCGTG
InFus-$MenT_3$K189A-left	GCAGTTACTGCGTGCAGCGCCTGCGCGATTTGCCAGGG
Cloning into pMPMK6/pK6	
$MenT_1$ EcoRI-For	TTGAATTCCATATGAACGCTGTGGAGTCGACACTCC
$MenT_1$ HindIII-Rev	TTAAGCTTGGATCCTTACCACTTGGCGGCGAGGC
$MenT_2$ EcoRI-For	TTGAATTCCATATGCTCGTCGGGGCACAGTGC
$MenT_2$ HindIII-Rev	TTAAGCTTGGATCCTTAGCCCGGTCGACCCACGG
$MenT_3$ MfeI-For	TTCAATTGCATATGACCAAGCCCTATTCGTC
$MenT_3$ HindIII-Rev	TTAAGCTTGGATCCTTATCTTTTCGTCGCCCGAT
$MenT_4$ MfeI-For	TTCAATTGCATATGGCCGGTCTGACCCGTGCGCTC
$MenT_4$ HindIII-Rev	TTAAGCTTGGATCCTTAGGACCGCAGCACCGCCA
Cloning into p29SEN	
$MenA_1$ EcoRI-For	TTGAATTCCATATGGCAGTTTCCGTCGCTGCGCAG
$MenA_1$ HindIII-Rev	TTAAGCTTGGATCCTTATGTGAACCGTGTGGACG
$MenA_2$ EcoRI-For	TTGAATTCCATATGGATCAGATCGGGGCTGA
$MenA_2$ HindIII-Rev	TTAAGCTTGGATCCTTATAATCGGGCGAGTCGCTC
$MenA_3$ EcoRI-For	TTGAATTCCATATGTTGTGTGCAAAACCGTATCT
$MenA_3$ HindIII-Rev	TTAAGCTTGGATCCTTACGCCGATGCTCGCTTCG
$MenA_4$ MfeI-For	TTCAATTGCATATGAATTCAAGTTCGGTGGTGAG
$MenA_4$ HindIII-Rev	TTAAGCTTGGATCCTTACGCCTTGCCGATCACGC
Rph EcoRI-For	TTGAATTCCATATGCGTCCAGCAGGCCGTAGC
Rph HindIII-Rev	TTAAGCTTGGATCCTCAGTTTGCCAGCGCCGCCT
Cloning into pGMC	

$MenT_1$ In-Fusion-For	GAAGACAGGCTGCCCATGAACGCTGTGGAGTCGAC
$MenT_1$ In-Fusion-Rev	TGTATAATAAAGTTGTTACCACTTGGCGGCGAGGC
$MenT_{1\text{-}His}$ In-Fusion-For	GAAGACAGGCTGCCCATGCACCACCACCACCACCACAGCAGCGGCAACGCTGTGG AGTCGACACT
$MenT_1$ In-Fusion-Rev	TGTATAATAAAGTTGTTACCACTTGGCGGCGAGGC
$MenT_2$ In-Fusion-For	GAAGACAGGCTGCCCATGCTCGTCGGGGCACAGTGC
$MenT_2$ In-Fusion-Rev	TGTATAATAAAGTTGTTAGCCCGGTCGACCCACGG
$MenT_3$ In-Fusion-For	GAAGACAGGCTGCCCATGACCAAGCCCTATTCGTC
$MenT_3$ In-Fusion-Rev	TGTATAATAAAGTTGTTATCTTTTCGTCGCCCGAT
$MenT_4$ In-Fusion-For	GAAGACAGGCTGCCCATGGCCGGTCTGACCCGTGC
$MenT_4$ In-Fusion-Rev	TGTATAATAAAGTTGTTAGGACCGCAGCACCGCCA
$MenA_4$-$MenT_4$ In-Fusion-For	GAAGACAGGCTGCCCATGAATTCAAGTTCGGTGGTGAG
$MenA_4$-$MenT_4$ In-Fusion-Rev	TGTATAATAAAGTTGTTAGGACCGCAGCACCGCCA

Cloning into pLAM

$MenA_1$ NdeI-For	TTGAATTCCATATGGCAGTTTCCGTCGCTGCGCAG
$MenA_1$ EcoRI-Rev	TTGAATTCTTATGTGAACCGTGTGGACG
$MenA_{1\text{-}His}$ NdeI-For	TTCATATGCACCACCACCACCACCACAGCAGCGGCGCAGTTTCCGTCGCTGCGCA
$MenA_1$ EcoRI-Rev	TTGAATTCTTATGTGAACCGTGTGGACG
$MenA_2$ NdeI-For	TTGAATTCCATATGGATCAGATCGGGGCTGA
$MenA_2$ EcoRI-Rev	TTGAATTCTTATAATCGGGCGAGTCGCTC
$MenA_3$ NdeI-For	TTGAATTCCATATGTTGTGTGCAAAACCGTATCT
$MenA_3$ EcoRI-Rev	TTGAATTCTTACGCCGATGCTCGCTTCG
$MenA_4$ NdeI-For	TTCAATTGCATATGAATTCAAGTTCGGTGGTGAG
$MenA_4$ MfeI-Rev	TTCAATTGTTACGCCTTGCCGATCACGC

Cloning into pET

$MenT_1$ NdeI-For	TTGAATTCCATATGAACGCTGTGGAGTCGACACTCC
$MenT_1$ MfeI-Rev	TTCAATTGTTACCACTTGGCGGCGAGGC
$MenT_{1\text{-}His}$ NdeI-For	TTCATATGCACCACCACCACCACCACAGCAGCGGCAACGCTGTGGAGTCGACACT
$MenT_{1\text{-}His}$ MfeI-Rev	TTCAATTGTTACCACTTGGCGGCGAGGC
$MenA_1$ NcoI-For	TTCCATGGCAGTTTCCGTCGCTGCGCA
$MenA_1$ BamHI-Rev	TTGGATCCTTATGTGAACCGTGTGGACG
$MenA_{1\text{-}His}$ NcoI-For	TTCCATGGTGCACCACCACCACCACCACAGCAGCGGCGCAGTTTCCGTCGCTGCGC A
$MenA_{1\text{-}His}$ BamHI-Rev	TTGGATCCTTATGTGAACCGTGTGGACG
$MenT_3$ NdeI-For	TTCAATTGCATATGACCAAGCCCTATTCGTC
$MenT_3$ MfeI-Rev	TTCATATGTATATCTCCTTCTTATACTT
$MenT_{3\text{-}His}$ NdeI-For	TTCAATTGCATATGACCAAGCCCTATTCGTC
$MenT_{3\text{-}His}$ MfeI-Rev	TTCAATTGAAGCTTTTAGTGGTGGTGGTGGTGGTGGCCGCTGCTTCTTTTCGTCGC CCGATCAA
$MenT_{3\text{-}His}$ D80A-For	TCGCGGACCTCCAAAGCCTTCGACACGGTCGCA
$MenT_{3\text{-}His}$ D80A-Rev	TGCGACCGTGTCGAAGGCTTTGGAGGTCCGCGA
$MenT_{3\text{-}His}$ K189A-For	TGGCAAATCGCGCAGGCGCTGCACGCAGTAACT
$MenT_{3\text{-}His}$ K189A-Rev	AGTTACTGCGTGCAGCGCCTGCGCGATTTGCCA
$MenA_3$ NcoI-For	TTCCATGGTGTTGTGTGCAAAACCGTATCT
$MenA_3$ HindIII-Rev	TTAAGCTTGGATCCTTACGCCGATGCTCGCTTCG
$MenA_{3\text{-}His}$ NcoI-For	TTCCATGGTGCACCACCACCACCACCACAGCAGCGGCTTGTGTGCAAAACCGTATC T
$MenA_{3\text{-}His}$ HindIII-Rev	TTAAGCTTGGATCCTTACGCCGATGCTCGCTTCG
$MenA_{3\text{-}His}$ NdeI-For	TTGAATTCCATATGTTGTGTGCAAAACCGTATCT
$MenA_{3\text{-}His}$ HindIII-Rev	TTAAGCTTGGATCCTTACGCCGATGCTCGCTTCG

Cloning into pTA100

$MenA_3$ MfeI-For (PF1330)	TTTCATATGCAATTGAGGAGGACAGGGATGTGTGCAAAACCGTATCTAA

$MenA_3$ XmaI-Rev (PF1331)	TTTACTAGTCCCGGGCTTGGTCACGCCGATG

Cloning into pBAD30

$MenT_3$ MfeI-For (PF1332)	TTTCATATGCAATTGAGGAGGACAGGGATGACCAAGCCCTATTCGTC
$MenT_3$ XmaI-Rev (PF1333)	TTTACTAGTCCCGGGTCATCTTTTCGTCGCCC
His_6-SUMO KpnI-For (TRB1120)	TTTGGTACCAAGAAGGAGATATATCCATGAGTGGC
His_6-SUMO-Rev (TRB1121)	GCCTCCCGTCTGCTGTTGAA
His_6-SUMO end/$MenT_3$ start-For (TRB1122)	TTCAACAGCAGACGGGAGGCACCAAGCCCTATTCGTCGCC
$MenT_3$ HindIII-Rev (TRB1124)	TTTAAGCTTTTATTATCTTTTCGTCGCCCGATCAA
His_6-SUMO XmaI-For (TRB1175)	TTTCCCGGGAAGAAGGAGATATATCCATGAGTGGC
His_6-SUMO end/ $MenT_4$ start-For (TRB1176)	TTCAACAGCAGACGGGAGGCGCCGGTCTGACCCGTGCG
$MenT_4$ HindIII-Rev (TRB1177)	TTTAAGCTTTTATTAGGACCGCAGCACCGCCAG
His_6-SUMO LIC site EcoRI-For (TRB1460)	TTTGAATTCTTTGTTTAACTTTAAGAAGGAGATATATCC
His_6-SUMO LIC site HindIII-Rev (TRB1462)	TTTAAGCTTGGATCCCTCGAGGTCGAC

Site-directed mutagenesis

$MenT_3$ K189A-For (TRB1525)	AATCGCGCAGGCACTGCACGCAG
$MenT_3$ K189A-Rev (TRB1526)	TGCCAGGGAATGGTCATG
$MenT_3$ D211A-For (TRB1674)	GACCTGGTGGCATTGCAGCTTCTTGAAG
$MenT_3$ D211A-Rev (TRB1675)	GTGAGCGCGGTCGTTGAC
$MenT_3$ D80A-For (TRB1676)	ACCTCCAAAGCATTCGACACGGTCG
$MenT_3$ D80A-Rev (TRB1677)	CCGCGAATCGGGAATTCC

Ligation independent cloning (LIC)

$MenA_3$ LIC-For	CAACAGCAGACGGGAGGTTGTGCAAAACCGTATCTAATTGATACGATTGCGC
$MenA_3$ LIC-Rev	GCGAGAACCAAGGAAAGGTTATTACGCCGATGCTCGCTTCGG
$MenT_3$ K189A LIC-For (TRB1020)	CAACAGCAGACGGGAGGTACCAAGCCCTATTCGTCGCCG
$MenT_3$ K189A LIC-Rev (TRB1021)	GCGAGAACCAAGGAAAGGTTATTATCTTTTCGTCGCCCGATCAATCCG

pMPMA-2 primers

pMPMA2-For	CGTAATACGACTCACTATAGG
pMPMA2-Rev	CCAAGCGCGCAATTAACCCTC

***in vitro* transcription/translation assay templates**

GatZ For	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATATGAAA ACGTTAATTGCCCGGCATAAA
GatZ Rev	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTCATTTTTCGAACTGCGGGTGGCT CCAGCTACCCCCTTCCGCACAGCCGTAGCGATA
WaaF For	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATATGAAA ATACTGGTGATCGGCCCGTCT
WaaF Rev	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTCATTTTTCGAACTGCGGGTGGCT CCAGCTACCCCCGGCTTCCTCTTGTAACAATAG

***E.coli* $tRNA^{Trp}$ DNA probe**

E.coli $tRNA^{Trp}$ DNA probe	CCCAACACCCGGTTTTGGAGACC

tRNA amplification

E. coli tRNA Trp For (CC2556)	ATTAATACGACTCACTATAGGGGCGTAGTTCAATTGGTAG
E. coli tRNA Trp Rev (CC2557)	TGGCAGGGGCGGAGAGACTC
E. coli tRNA Trp Rev no CCA motif (CC2591)	CAGGGGCGGAGAGACTCGAAC
M. tuberculosis tRNA Ala-1 For	ATTAATACGACTCACTATAGGGGGCTATGGCGCAGTTGGTAGCGCG
M. tuberculosis tRNA Ala-1 Rev	TGGTGGAGCTAAGGGGATTC

M. tuberculosis tRNA Ala-2 For	ATTAATACGACTCACTATAGGGGGCTATGGCGCAGCTGGTAGCGCA
M. tuberculosis tRNA Ala-2 Rev	TGGTGGAGCTAAGGGGACTC
M. tuberculosis tRNA Ala-3 For	ATTAATACGACTCACTATAGGGGGCCTTAGCTCAGTTGGTAGA
M. tuberculosis tRNA Ala-3 Rev	TGGTGGAGCCTAGGGGACTC
M. tuberculosis tRNA Arg-1 For	ATTAATACGACTCACTATAGGCGCCCGTAGCTCAACGGATAGAGCATC
M. tuberculosis tRNA Arg-1 Rev	TGGCGCGCCCGAAGAGATTC
M. tuberculosis tRNA Arg-2 For	ATTAATACGACTCACTATAGGCCCCCGTAGCTCAGGGGATAGAGCG
M. tuberculosis tRNA Arg-2/3 Rev	TGGTGCCCCCGGCAGGATTC
M. tuberculosis tRNA Arg-3 For	ATTAATACGACTCACTATAGGCCCTCGTAGCTCAGGGGATAGAGCACG
M. tuberculosis tRNA Arg-2/3 Rev	TGGTGCCCCCGGCAGGATTC
M. tuberculosis tRNA Arg-4 For	ATTAATACGACTCACTATAGGCCTCCGTAGCTCAGGTGGA
M. tuberculosis tRNA Arg-4 Rev	TGGTGCCCCCGGCACGACTC
M. tuberculosis tRNA Asn For	ATTAATACGACTCACTATAGTCCCCTGTAGCTCAATTGGCA
M. tuberculosis tRNA Asn Rev	TGGCTCCCCCGGGAGGACTC
M. tuberculosis tRNA Asp For	ATTAATACGACTCACTATAGGGCCCTGTGGCGCAGTTGGT
M. tuberculosis tRNA Asp Rev	TGGCGACCCTGACGGGACTCG
M. tuberculosis tRNA Cys For	ATTAATACGACTCACTATAGGGTGGAGTGGCCGAGTGGTG
M. tuberculosis tRNA Cys Rev	TGGAGGTGGAGACGGGAATC
M. tuberculosis tRNA Gln-1 For	ATTAATACGACTCACTATAGTGGGGTATGGTGTAATTGGCAA
M. tuberculosis tRNA Gln-1 Rev	TGGCTGGGGTACCAGGACTC
M. tuberculosis tRNA Gln-2 For	ATTAATACGACTCACTATAGTCCGTCGTGGTGTAATCGGCAG
M. tuberculosis tRNA Gln-2 Rev	TGGCTCCGTCGCCAGGACTC
M. tuberculosis tRNA Glu-1 For	ATTAATACGACTCACTATAGGGCCCCGTCGTCTAGCGGCCTA
M. tuberculosis tRNA Glu-1 Rev	TGGTAGCCCCGATGGGATTC
M. tuberculosis tRNA Glu-2 For	ATTAATACGACTCACTATAGGCCCCCTTCGTCTAGACGGCCT
M. tuberculosis tRNA Glu-2 Rev	TGGTACCCCCTACGGGATTC
M. tuberculosis tRNA Gly-1 For	ATTAATACGACTCACTATAGGCCGATGTAGTTCAATGGC
M. tuberculosis tRNA Gly-1 Rev	TGGAGCCGATGACGGGAATC
M. tuberculosis tRNA Gly-2 For	ATTAATACGACTCACTATAGGCGGATGTAGCGCAGTTGGT
M. tuberculosis tRNA Gly-2 Rev	TGGAGCGGATGACGGGATTC
M. tuberculosis tRNA Gly-3 For	ATTAATACGACTCACTATAGGCGGGCGTAGCTCAATGGT
M. tuberculosis tRNA Gly-3 Rev	TGGAGCGGGCGACGGGAATC
M. tuberculosis tRNA His For	ATTAATACGACTCACTATAGGTGAGTGTAGTTCAGTTGGT
M. tuberculosis tRNA His Rev	TGGGGTGAGTGACGGGACTC
M. tuberculosis tRNA Ile For	ATTAATACGACTCACTATAGGGGCCTATAGCTCAGGCGGT
M. tuberculosis tRNA Ile Rev	TGGTGGGCCTAGGAGGACTCGAA
M. tuberculosis tRNA Leu-1 For	ATTAATACGACTCACTATAGGCCCTCGTATCCCAACTGGCAGAGGA
M. tuberculosis tRNA Leu-1 Rev	TGGTGCCCTCGGTGGGATTC
M. tuberculosis tRNA Leu-2 For	ATTAATACGACTCACTATAGGGGCGAGTGGCGGAATGGCAGACGCGCTG
M. tuberculosis tRNA Leu-2 Rev	TGGTGGGCGAAGGGGGACTT
M. tuberculosis tRNA Leu-3 For	ATTAATACGACTCACTATAGGTCCGAGTGGCGGAATGGCAGACGCGCTA
M. tuberculosis tRNA Leu-3 Rev	TGGTGTCCGAGGGGGGACTT
M. tuberculosis tRNA Leu-4 For	ATTAATACGACTCACTATAGGCCCCCATAGCCCAATTGGCAGAGGC
M. tuberculosis tRNA Leu-4 Rev	TGGTGCCCCCAGTCGGACTC
M. tuberculosis tRNA Leu-5 For	ATTAATACGACTCACTATAGGCGGGCGTGATGAAATTGGCAA
M. tuberculosis tRNA Leu-5 Rev	TGGTGCGGGCGGAGGGACTC
M. tuberculosis tRNA Lys-1 For	ATTAATACGACTCACTATAGGCGCCGTTAGCTCAGTTGGTAGAGCAG
M. tuberculosis tRNA Lys-1 Rev	TGGTGCGCCGTCAGGGTTTCGAA
M. tuberculosis tRNA Lys-2 For	ATTAATACGACTCACTATAGGCCCCTATAGCTCAGTTGGTAGAGCTA
M. tuberculosis tRNA Lys-2 Rev	TGGTGCCCCCACCAGGACTCGAA
M. tuberculosis tRNA Met-1 For	ATTAATACGACTCACTATAGCGCGGGGTGGAGCAGCTCGG
M. tuberculosis tRNA Met-1 Rev	TGGAGCGGGGACAGGATTCGAAC
M. tuberculosis tRNA Met-2 For	ATTAATACGACTCACTATAGGGCGATGTAGCTCAGTCGGTTAGAGCGA
M. tuberculosis tRNA Met-2 Rev	TGGTAGCGATGGCCGGACTCGAA
M. tuberculosis tRNA Met-3 For	ATTAATACGACTCACTATAGGGGGCGGTAGCTCAGTTGGTTAGAGCCG
M. tuberculosis tRNA Met-3 Rev	TGGTGGGGCGGGCGGGGCTCGAA
M. tuberculosis tRNA Phe For	ATTAATACGACTCACTATAGGGCCAGGTAGCTCAGTCGGT

M. tuberculosis tRNA Phe Rev	TGGTGGCCAGGGGCGGGATC
M. tuberculosis tRNA Pro-1 For	ATTAATACGACTCACTATAGCGGGGTGTGGCGCAGCTTGGTAGCGCG
M. tuberculosis tRNA Pro-1 Rev	TGGTCGGGGTGGCGGGATTTGAA
M. tuberculosis tRNA Pro-2 For	ATTAATACGACTCACTATAGCGGGCTGTGGCGCAGTTTGGTAGCGCAC
M. tuberculosis tRNA Pro-2 Rev	TGGTCGGGCTGACAGGATTTGAA
M. tuberculosis tRNA Pro-3 For	ATTAATACGACTCACTATAGCGGGGTGTAGCGCAGCTTGGTAGCGCAT
M. tuberculosis tRNA Pro-3 Rev	TGGTCGGGGTGACAGGATTT
M. tuberculosis tRNA Ser-1 For	ATTAATACGACTCACTATAGGGTGGCGTGTCCGAGCGGCCTAAG
M. tuberculosis tRNA Ser-1/4 Rev	TGGCGGTGGCGGAGGGATTT
M. tuberculosis tRNA Ser-2 For	ATTAATACGACTCACTATAGGGAGGCGTGCCAGAGCGGCCGA
M. tuberculosis tRNA Ser-2 Rev	TGGCGGAGGCGAGAGGATTT
M. tuberculosis tRNA Ser-3 For	ATTAATACGACTCACTATAGGGAGGATTCGCCTAGTGGCCTAT
M. tuberculosis tRNA Ser-3 Rev	TGGCGGAGGATGCGGGATTT
M. tuberculosis tRNA Ser-4 For	ATTAATACGACTCACTATAGGGTGGCGTGGCAGAGCGGCCTAAT
M. tuberculosis tRNA Ser-1/4 Rev	TGGCGGTGGCGGAGGGATTT
M. tuberculosis tRNA Thr-1 For	ATTAATACGACTCACTATAGGCCGCCTTAGCTCAGTCGGT
M. tuberculosis tRNA Thr-1 Rev	TGGAGCCGCCTAGGAGAATC
M. tuberculosis tRNA Thr-2 For	ATTAATACGACTCACTATAGGCCCCCTTAGCTCAGTCGGC
M. tuberculosis tRNA Thr-2 Rev	TGGAGCCCCCTAACGGAATC
M. tuberculosis tRNA Thr-3 For	ATTAATACGACTCACTATAGGCCTCCTTAGCTCAGTGGTA
M. tuberculosis tRNA Thr-3 Rev	TGGAGCCCCCTGTCAGGATT
M. tuberculosis tRNA Trp For	ATTAATACGACTCACTATAGAGGGGCGTAGCTCAACTGGC
M. tuberculosis tRNA Trp Rev	TGGCAGGGGCGACAGGACTTGAA
M. tuberculosis tRNA Tyr For	ATTAATACGACTCACTATAGGGCAGGTTGCCCGAGCGGCC
M. tuberculosis tRNA Tyr Rev	TGGTGGCAGGTGCAGGATTC
M. tuberculosis tRNA Val-1 For	ATTAATACGACTCACTATAGGGTCCCGTGGCTCAGTGGGAGAGCGT
M. tuberculosis tRNA Val-1 Rev	TGGTGGTCCCGGCTGGGATC
M. tuberculosis tRNA Val-2 For	ATTAATACGACTCACTATAGGCGCGATTAGCTCAGCGGGAGAGCGC
M. tuberculosis tRNA Val-2 Rev	TGGTGCGCGATACTGGGATT
M. tuberculosis tRNA Val-3 For	ATTAATACGACTCACTATAGGGGCGCGTAGCTCAGCGGT
M. tuberculosis tRNA Val-3 Rev	TGGTGGGCGCGGACGGGATC

Table S2. Plasmids used in this study

Plasmid	Parent vector	Cloning technique	Primer set/Restriction enzymes used	Reference
p29SEN	Ap^R	-	-	(*40*)
p29SEN-$MenA_1$	p29SEN	Restriction site cloning	EcoRI/HindIII	This work
p29SEN-$MenA_2$	p29SEN	Restriction site cloning	EcoRI/HindIII	This work
p29SEN-$MenA_3$	p29SEN	Restriction site cloning	EcoRI/HindIII	This work
p29SEN-$MenA_4$	p29SEN	Restriction site cloning	MfeI/HindIII	This work
p29SEN-Rph	p29SEN	Restriction site cloning	EcoRI/HindIII	This work
pBAD30	Ap^R	-	-	(*44*)
pDE43-MCS	Sm^R	-	-	(*18*)
pETDuet-1	Ap^R	-	-	Novagen
pET-$MenT_1$	pETDuet-1	Restriction site cloning	NdeI/MfeI	This work
pET-$MenT_{1\text{-}His}$	pETDuet-1	Restriction site cloning	NdeI/MfeI	This work
pET-$MenT_{1\text{-}His}$/$MenA_1$	pET-$MenT_{1\text{-}His}$	Restriction site cloning	NcoI/BamHI	This work
pET-$MenA_1$	pETDuet-1	Restriction site cloning	NcoI/BamHI	This work
pET-$MenA_{1\text{-}His}$	pETDuet-1	Restriction site cloning	NcoI/BamHI	This work
pET-$MenT_1$/$MenA_{1\text{-}His}$	pET-$MenA_{1\text{-}His}$	Restriction site cloning	NdeI/MfeI	This work
pET-$MenT_3$	pETDuet-1	Restriction site cloning	NdeI/MfeI	This work
pET-$MenT_{3\text{-}His}$	pETDuet-1	Restriction site cloning	NdeI/MfeI	This work
pET-$MenT_{3\text{-}His}$/$MenA_3$	pET-$MenT_{3\text{-}His}$	Restriction site cloning	NcoI/HindIII	This work
pET-$MenT_{3\text{-}His}^{(D80A)}$	pETDuet-1	Restriction site cloning	$MenT_{3\text{-}His}$ D80A-For $MenT_{3\text{-}His}$ D80A-Rev	This work
pET-$MenT_{3\text{-}His}^{(K189A)}$	pETDuet-1	Restriction site cloning	$MenT_{3\text{-}His}$ K189A-For $MenT_{3\text{-}His}$ K189A-Rev	This work
pET-$MenA_3$	pETDuet-1	Restriction site cloning	NcoI/HindIII	This work
pET-$MenA_{3\text{-}His}$	pETDuet-1	Restriction site cloning	NcoI/HindIII	This work
pET-$MenT_3$/$MenA_{3\text{-}His}$	pET-$MenA_{3\text{-}His}$	Restriction site cloning	NdeI/MfeI	This work
pET15b	Ap^R	-	-	Novagen
pET-$MenA_{3\text{-}His}$	pET15b	Restriction site cloning	NdeI/HindIII	This work
pGMCS	Sm^R	-	-	(*41*)

pGMC-$MenT_1$	pGMCS	In-Fusion cloning	$MenT_1$ In-Fusion-For $MenT_1$ In-Fusion-Rev	This work
pGMC-$MenT_{1\text{-His}}$	pGMCS	In-Fusion cloning	$MenT_{1\text{-His}}$ In-Fusion-For $MenT_1$ In-Fusion-Rev	This work
pGMC-$MenT_2$	pGMCS	In-Fusion cloning	$MenT_2$ In-Fusion-For $MenT_2$ In-Fusion-Rev	This work
pGMC-$MenT_3$	pGMCS	In-Fusion cloning	$MenT_3$ In-Fusion-For $MenT_3$ In-Fusion-Rev	This work
pGMC-$MenT_4$	pGMCS	In-Fusion cloning	$MenT_4$ In-Fusion-For $MenT_4$ In-Fusion-Rev	This work
pGMC-*$MenA_4$-$MenT_4$	pGMCS	In-Fusion cloning	$MenA_4$ In-Fusion-For $MenT_4$ In-Fusion-Rev	This work
pGMCS-P1-$MenA_3$	pDE43-MCS	Multisite gateway	clo-RBS1-$MenA_3$-attB2/clo-$MenA_3$-attB3	This work
pGMCS-TetR-P1-RBS1-$MenA_3$	pDE43-MCS	Multisite gateway	clo-RBS1-$MenA_3$-attB2/clo-$MenA_3$-attB3	This work
pGMCS-TetR-P1-RBS1-$MenA_3$-$MenT_3$	pDE43-MCS	Multisite gateway	clo-RBS1-$MenA_3$-attB2/clo-$MenT_3$-attB3	This work
pGMCS-TetR-P1-RBS1-$MenT_3$	pDE43-MCS	Multisite gateway	clo-RBS1-$MenT_3$-attB2/clo-$MenT_3$-attB3	This work
pGMCS-TetR-P1-RBS4-$MenT_3$	pDE43-MCS	Multisite gateway	clo-RBS4-$MenT_3$-attB2/clo-$MenT_3$-attB3	This work
pGMCS-TetR-P1-RBS4-$MenT_3$(D80A)	pGMCS-TetR-P1-RBS4-$MenT_3$	In-Fusion cloning	InFus-$MenT_3$D80A-right InFus-$MenT_3$D80A-left	This work
pGMCS-TetR-P1-RBS4-$MenT_3$(K189A)	pGMCS-TetR-P1-RBS4-$MenT_3$	In-Fusion cloning	InFus-$MenT_3$K189A-right InFus-$MenT_3$K189A-left	This work
pGMCS-TetR-P1-RBS4-$MenT_3$(D211A)	pGMCS-TetR-P1-RBS4-$MenT_3$	In-Fusion cloning	InFus-$MenT_3$D211A-right InFus-$MenT_3$D211A-left	This work
pGMCZ	Zc^R	-	-	(*42*)
pJV53H	pJV53; Hm^R	-	-	(*43*)
pLAM12	Km^R	-	-	(*43*)
pLAM-$MenA_1$	pLAM12	Restriction site cloning	NdeI/EcoRI	This work
pLAM-$MenA_{1\text{-His}}$	pLAM12	Restriction site cloning	NdeI/EcoRI	This work
pLAM-$MenA_2$	pLAM12	Restriction site cloning	NdeI/EcoRI	This work
pLAM-$MenA_3$	pLAM12	Restriction site cloning	NdeI/EcoRI	This work
pLAM-$MenA_4$	pLAM12	Restriction site cloning	NdeI/MfeI	This work
pMPMA2-based multicopy plasmid library of *E. coli* genes	Ap^R	-	-	(*48*)

pMPMK6	Km^R	-	-	(*39*)
pK6-$MenT_1$	pMPMK6	Restriction site cloning	EcoRI/HindIII	This work
pK6-$MenT_2$	pMPMK6	Restriction site cloning	EcoRI/HindIII	This work
pK6-$MenT_3$	pMPMK6	Restriction site cloning	MfeI/HindIII	This work
pK6-$MenT_4$	pMPMK6	Restriction site cloning	MfeI/HindIII	This work
pPF656	pTA100	Restriction site cloning	PF1330/PF1331	This work
pPF657	pBAD30	Restriction site cloning	PF1332/PF1333	This work
pRARE	Cm^R	-	-	Novagen
pSAT1-LIC	Ap^R	-	-	This work
pTA100	Sm^R	-	-	(*4*)
pTRB491	pSAT1-LIC	Ligation Independent Cloning	TRB1018/TRB1019	This work
pTRB517	pBAD30	Restriction site cloning	TRB1120/TRB1121/TRB1122/TRB1124	This work
pTRB544	pBAD30	Restriction site cloning	TRB1121/TRB1175/TRB1176/TRB1177	This work
pTRB550	pBAD30	Restriction site cloning	TRB1460/TRB1462	This work
pTRB562	pPF657	Inverse PCR site-directed mutagenesis	TRB1525/TRB1526	This work
pTRB577	pTRB550	Ligation independent cloning	TRB1020/TRB1021	This work
pTRB591	pPF657	Inverse PCR site-directed mutagenesis	TRB1676/TRB1677	This work
pTRB592	pPF657	Inverse PCR site-directed mutagenesis	TRB1674/TRB1675	This work
pTRB593	pTRB517	Inverse PCR site-directed mutagenesis	TRB1676/1677	This work

Journal of Structural Biology 213 (2021) 107752

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structure Report

Crystal structure of the anti-CRISPR repressor Aca2

Ben Usher [a], Nils Birkholz [b], Izaak N. Beck [a], Robert D. Fagerlund [b,c], Simon A. Jackson [b,c], Peter C. Fineran [b,c,d], Tim R. Blower [a,*]

[a] Department of Biosciences, Durham University, Stockton Road, Durham DH1 3LE, UK
[b] Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin 9054, New Zealand
[c] Genetics Otago, University of Otago, PO Box 56, Dunedin 9054, New Zealand
[d] Bio-Protection Research Centre, University of Otago, PO Box 56, Dunedin 9054, New Zealand

ARTICLE INFO

Edited by H. Engel Andreas

Keywords:
CRISPR
X-ray crystallography
Anti-CRISPR associated
Transcriptional regulator
Aca2

ABSTRACT

Bacteria use adaptive CRISPR-Cas immune mechanisms to protect from invasion by bacteriophages and other mobile genetic elements. In response, bacteriophages and mobile genetic elements have co-evolved anti-CRISPR proteins to inhibit the bacterial defense. We and others have previously shown that anti-CRISPR associated (Aca) proteins can regulate this anti-CRISPR counter-attack. Here, we report the first structure of an Aca protein, the Aca2 DNA-binding transcriptional autorepressor from *Pectobacterium carotovorum* bacteriophage ZF40, determined to 1.34 Å. Aca2 presents a conserved N-terminal helix-turn-helix DNA-binding domain and a previously uncharacterized C-terminal dimerization domain. Dimerization positions the Aca2 recognition helices for insertion into the major grooves of target DNA, supporting its role in regulating anti-CRISPRs. Furthermore, database comparisons identified uncharacterized Aca2 structural homologs in pathogenic bacteria, suggesting that Aca2 represents the first characterized member of a more widespread family of transcriptional regulators.

1. Introduction

Bacteria are under constant threat of invasion by bacteriophages (phages) and other mobile genetic elements (MGEs). Among the many protection strategies employed against these invaders, the highly diverse CRISPR-Cas systems stand out as the only known adaptive immune systems in bacteria (Hampton et al., 2020). In response, phages and MGEs have evolved a large array of anti-CRISPR (Acr) proteins which can inhibit CRISPR-Cas defense through various means (Malone et al., 2021; Wiegand et al., 2020). With different Cas proteins, such as Cas9, being utilized as tools in bioengineering, Acr proteins offer a way to make these tools more controllable and may substantially facilitate their application (Marino et al., 2020).

Many anti-CRISPR genes form an operon with genes encoding anti-CRISPR-associated (Aca) proteins, ten families of which have been identified (Bondy-Denomy et al., 2013; He et al., 2018; León et al., 2021; Marino et al., 2018; Pawluk et al., 2016a, 2016b; Pinilla-Redondo et al., 2020; Yin et al., 2019). For example, *Pseudomonas aeruginosa* phage JBD30 contains an *acrIF1–aca1* operon and *Pectobacterium carotovorum* phage ZF40 contains an *acrIF8–aca2* operon (Fig. 1A). We and others recently showed that Aca1 and Aca2, as well as Aca3 encoded in an *acrIIC3–aca3* operon, serve as repressors of their respective promoters (Birkholz et al., 2019; Stanley et al., 2019). These findings and the pervasive presence of helix-turn-helix (HTH) domains in all known Aca proteins suggest that Aca proteins generally function to repress, or at least to regulate, anti-CRISPR production. In some cases, the anti-CRISPR itself contains an HTH domain for autoregulation (Osuna et al., 2020). Interestingly, bacteria may use their own Aca-like regulators to inhibit anti-CRISPR deployment by phages, thereby maintaining CRISPR-Cas defense (Osuna et al., 2020; Stanley et al., 2019).

Aca1 and Aca2 bind to inverted repeats (IRs) that overlap with the −10 and −35 elements of their respective promoters, suggesting that transcriptional repression occurs through blocking of RNA polymerase recruitment (Birkholz et al., 2019; Stanley et al., 2019). The *acrIF8–aca2* promoter contains two similar IR pairs and IR1 was shown to be bound tightly by Aca2 (Fig. 1A). We demonstrated this interaction involves DNA bending, thus providing a first insight into the topological changes involved in anti-CRISPR regulation (Birkholz et al., 2019). However, information on the structural basis of Aca-mediated repression is still missing. In this study, we determined the crystal structure of the anti-CRISPR-associated protein Aca2 from *P. carotovorum* phage ZF40 to better understand its role as a transcriptional regulator.

* Corresponding author.
E-mail address: timothy.blower@durham.ac.uk (T.R. Blower).

https://doi.org/10.1016/j.jsb.2021.107752
Received 23 April 2021; Received in revised form 26 May 2021; Accepted 6 June 2021
Available online 8 June 2021

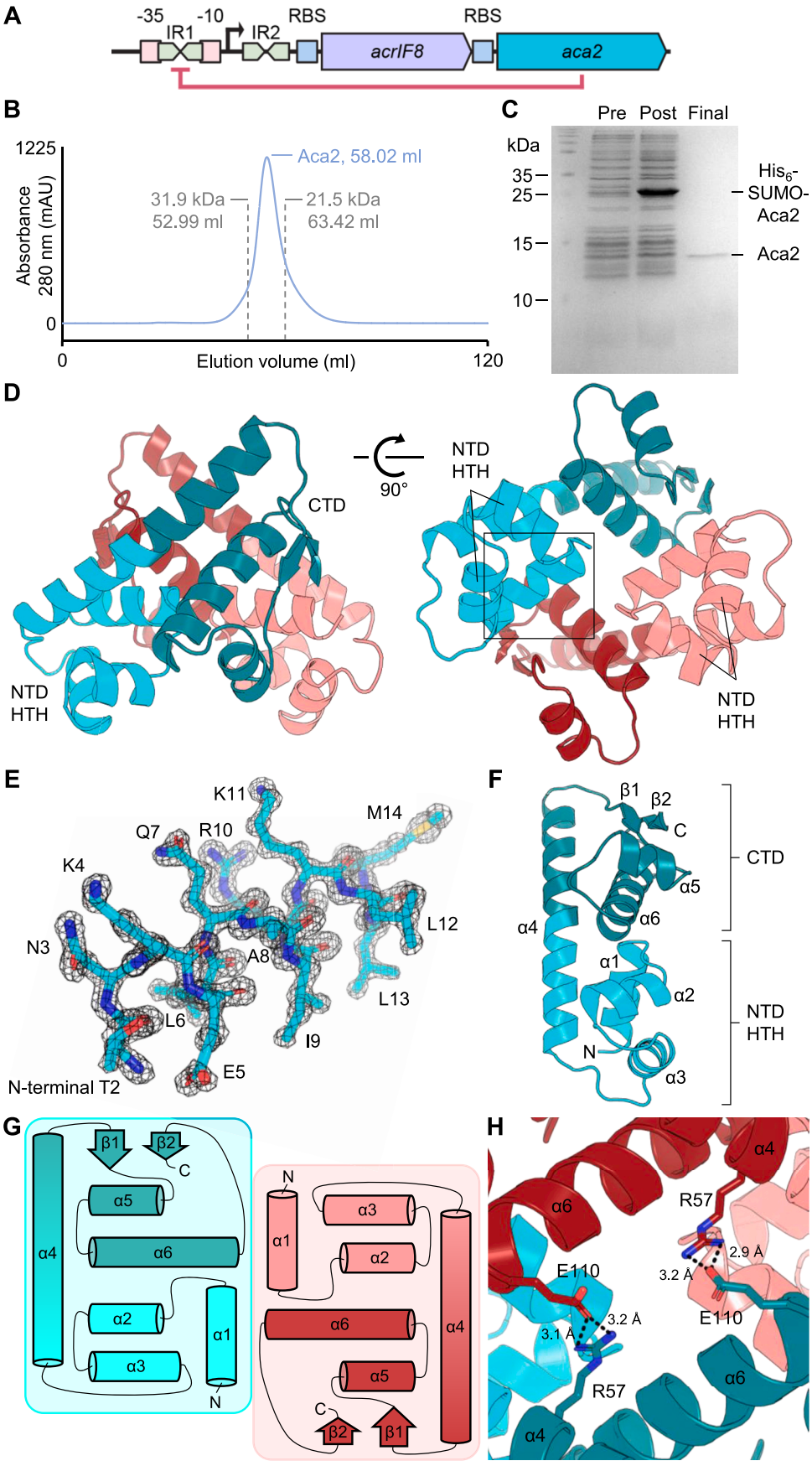


Fig. 1. Structure of Aca2. (A) Architecture of the *acrIF8–aca2* locus from *P. carotovorum* phage ZF40 (not to scale). Promoter elements (−10 and −35 regions), inverted repeats (IR1 and IR2), transcription start site (arrow) and ribosome-binding sites (RBS) are indicated. Aca2 binding to IR1 and repressing transcription is shown in red. (B) Elution volume of untagged Aca2 during size-exclusion chromatography (SEC) shows it is a dimer in solution. (C) SDS-PAGE of pre-induction (Pre), post-induction (Post), and cleaved, purified Aca2 protein (Final). (D) Cartoon overview of the Aca2 dimer, with one protomer shown as cyan (NTD) and teal (CTD), and the other protomer shown as pink (NTD) and red (CTD). Two orthogonal views are shown, rotated by 90°. NTD HTH – N-Terminal Domain Helix-Turn-Helix. CTD – C-Terminal Domain. (E) Boxed region of (D), containing helix α1 as sticks, shown with a 2Fo-Fc electron density map contoured to 2σ. (F) Single protomer of Aca2 with secondary structures and domains labelled. (G) Topology of the Aca2 dimer. (H) Close-up top view of the Aca2 dimer, rotated down by 90° from (D, left panel) showing salt bridges between protomers. Distances shown are in angstroms.

2. Materials and methods

2.1. Aca2 cloning

The *aca2* gene was amplified from pPF1575 (Birkholz et al., 2019) with primers TRB1765 (5′-CAACAGCAGACGGGAGGTACAAACAAA-GAACTTCAGGC-3′) and TRB1766 (5′-GCGAGAACCAAGGAAAGG TTATTATTAGATTAAATCCGCGTGACC-3′), then cloned into pSAT1-LIC (Cai et al., 2020) via ligation-independent cloning (LIC) (Aslanidis and de Jong, 1990), to produce pTRB627. The pSAT1-LIC plasmid features a LIC site that fuses an N-terminal His_6-SUMO tag to the target protein.

2.2. Recombinant protein expression

Aca2 was expressed in *E. coli* BL21(λDE3)Δ*slyD* (Cai et al., 2020) transformed with pTRB627. Overnight cultures were re-seeded 1:100 into 2 L baffled flasks containing 1 L 2×YT. Cells were grown at 160 rpm, 37 °C, until an OD_{600} of 0.3 was reached and then at 25 °C until OD_{600} 0.6. Expression was induced by the addition of IPTG (1 mM), then cells were left to grow overnight at 16 °C, with shaking at 160 rpm.

2.3. Recombinant protein purification

Following overnight expression, bacteria were harvested by centrifugation at 4,200 *g*, 4 °C, and the pellets were resuspended in buffer A [20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 5 mM imidazole, and 10% glycerol]. Cells were lysed by sonication at 40 kpsi and then centrifuged at 45,000 *g*, 4 °C. The clarified lysate was then passed down a HisTrap HP column (Cytiva) using a peristaltic pump. The resin-bound protein was first washed for 10 column volumes with buffer A, followed by 10 column volumes of buffer B [20 mM Tris-HCl (pH 7.9), 100 mM NaCl, 35 mM imidazole, and 10% glycerol] and then eluted directly onto a HiTrap Q HP column (Cytiva) with buffer C [20 mM Tris-HCl (pH 7.9), 100 mM NaCl, 250 mM imidazole, and 10% glycerol]. The Q HP column was washed briefly with 5 column volumes of buffer D [20 mM Tris-HCl (pH 7.9), 100 mM NaCl, 5 mM imidazole, and 10% glycerol], and then transferred to an Äkta Pure (Cytiva). Proteins were separated using an elution gradient from 100% buffer D to 40% buffer E [20 mM Tris-HCl (pH 7.9), 1 M NaCl, and 10% glycerol]. Fractions corresponding to the chromatogram protein peak were pooled and incubated overnight at 4 °C with hSENP2 SUMO protease to cleave the N-terminus His_6-SUMO tag from recombinant Aca2. The next day, the sample was passed through a second HisTrap HP column via a peristaltic pump, then washed for 2 column volumes with buffer A. The flow-through and wash fractions containing untagged Aca2 were collected and concentrated, then loaded onto a HiPrep 16/60 Sephacryl S-200 size exclusion column (Cytiva) connected to an Äkta Pure, in buffer S [50 mM Tris-HCl (pH 7.9), 500 mM KCl, and 10% glycerol]. Fractions corresponding to the chromatogram peak were analyzed by SDS-PAGE, with optimal fractions then pooled and dialyzed overnight at 4 °C into buffer X [20 mM Tris-HCl (pH 7.9), 150 mM NaCl, and 2.5 mM dithiothreitol (DTT)] for crystallography. Crystallography samples were concentrated, quantified, and stored on ice, then either used immediately or flash-frozen in liquid N_2 for storage at −80 °C.

2.4. Protein crystallization

Aca2 was concentrated to 12 mg ml^{-1} in buffer X. Crystallization screens were performed using a Mosquito Xtal3 robot (STP Labtech) to set 200:100 nl and 100:100 nl protein:condition sitting drops. Initial crystals formed 6 days after incubation at 294 K and were left to grow until day 29. Aca2 was observed to form small needle-like crystals in condition B6 of Structure 1 + 2 Eco (Molecular Dimensions) [0.2 M sodium acetate trihydrate, 100 mM MES (pH 6.5), and 30% w/v PEG 8000], at a final protein concentration of 8 mg ml^{-1}. The Aca2 crystals were harvested directly from the screen. To harvest, 20 µl of the condition reservoir was added to 20 µl of cryo buffer [25 mM Tris-HCl (pH 7.9), 187.5 mM NaCl, 3.125 mM DTT, and 80% glycerol] and mixed quickly by vortexing. An equal volume of this mixture was then added directly to the crystal drop, and the Aca2 crystal was immediately extracted using a nylon loop and flash-frozen in liquid N_2.

2.5. Data collection and structure determination

Diffraction data were recorded at 100 K on beamline I04 at Diamond Light Source. A single 360° dataset was collected for Aca2. Diffraction data were processed with XDS (Kabsch, 2010), and then AIMLESS in CCP4 (Winn et al., 2011) was used to corroborate the space group. The crystal structure of Aca2 was solved *ab initio* using ARCIMBOLDO (Rodríguez et al., 2009), with initial model-building then performed using Buccaneer (Cowtan, 2006) in CCP4 (Winn et al., 2011). Data processing then moved to PHENIX (Adams et al., 2010) and COOT (Emsley and Cowtan, 2004), where the model was iteratively refined and built, respectively. The quality of the final model was assessed using COOT and the wwPDB validation server (Gore et al., 2012). Structural figures were generated using PyMol (Schrödinger). RMSD values were calculated with the Super command in PyMol, using all atoms and then rejecting outlier pairs.

3. Results

3.1. Overall structure of Aca2

Aca2 was expressed and purified as described (Materials and Methods). The elution volume during the final size exclusion chromatography run indicated that the 13.7 kDa Aca2 protein forms a dimer in solution (Fig. 1B), corroborating what was observed with our previous constructs (Birkholz et al., 2019). This final Aca2 product was judged sufficiently pure for crystallization by SDS-PAGE (Fig. 1C). Using this sample, we determined the crystal structure of Aca2 to 1.34 Å (Fig. 1D) and refined the structure to an R-factor of 0.1476 and an R-free of 0.1761 (Table 1).

Previous work identified a putative N-terminal domain (NTD) containing an HTH motif required for DNA binding (Birkholz et al., 2019). The Aca2 dimer structure shows each protomer stacked against one another in opposition, like the letter X, such that the HTH motifs are aligned along the “base” of the dimer (Fig. 1D). The obtained data (Table 1) allowed all amino acids within an Aca2 dimer to be modelled (Fig. 1D), and an example section of the Aca2 2Fo-Fc electron density map is shown for the first alpha-helical region, beginning with the N-terminal amino acid T2 (Fig. 1E). Examining a single protomer shows that the proposed NTD HTH and relative C-terminal domain (CTD) are in fact small clusters of secondary structure elements abutting and joined by a longer backbone α-helix, α4, such that the protomer forms a single globular protein (Fig. 1F). All Aca2 residues (116 amino acids in total) are resolved in the structure except the initial methionine, which was not included in the construct. Aca2 is comprised of 6 α-helices; α1 (amino acid (aa) positions 2–13), α2 (aa 16–24), α3 (aa 28–38), α4 (aa 43–70), α5 (aa 81–89) and α6 (aa 93–110). A short β-strand, β1 (aa 74–78), is encoded between α4 and α5, and forms a very short two-stranded parallel β-sheet with β2 (aa 114–116) (Fig. 1F). An HTH motif contains an α-helix for positioning, and an α-helix for DNA recognition, linked by a short turn. In Aca2, α2 will help position α3 for DNA recognition (Fig. 1F). This is further supported by a previous mutagenesis study that showed R30 was necessary for promoter autoregulation by Aca2, and R30 is found on α3 (Birkholz et al., 2019). Whilst the NTD provides the HTH motif for DNA-binding, the CTD stabilizes the positioning of the NTD from the other protomer by stacking α6 against the other protomer α1, thereby aiding dimerization. This interaction forms the bulk of the dimer interface and can be seen both in the provided views (Fig. 1D), as well as schematically within the topology diagram (Fig. 1G). The Aca2 protomer-protomer interface was analyzed

Table 1
Data collection and refinement statistics for Aca2.

PDB ID code	75BJ
Data Collection	
Beamline	Diamond I04
Wavelength (Å)	0.9795
Resolution range (Å)[a]	38.19–1.34 (1.388–1.34)
Space group	$P2_1$
Unit cell, *a b c* (Å); α β γ (°)	39.791, 67.103, 42.240; 90, 106.331, 90
Total reflections[a]	646,247 (63250)
Unique reflections[a]	47,804 (4740)
Multiplicity[a]	13.5 (13.3)
Completeness (%)[a]	99.98 (99.96)
Mean $I/\sigma(I)$[a]	10.28 (1.18)
R_{merge}[a,b]	0.1588 (1.688)
$CC_{1/2}$[a]	0.999 (0.663)
Refinement	
R_{work}[c]	0.1476 (0.2103)
R_{free}[c]	0.1761 (0.2509)
Number of non-hydrogen atoms	2125
macromolecules	1854
ligands	6
solvent	265
Protein residues	230
RMS (bonds, Å)	0.009
RMS (angles, °)	1.07
Ramachandran favored (%)	99.12
Ramachandran allowed (%)	0.88
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	0.82
Average B-factor	19.31
macromolecules	17.48
ligands	24.20
solvent	32.04

[a] Statistics for the highest resolution shell are shown in parentheses.
[b] $R_{\text{merge}} = \Sigma_h\Sigma_i|I_{h,i}-I_h|/\Sigma_h\Sigma_i I_{h,i}$, where I_h is the mean intensity of the i observations of symmetry related reflections of h.
[c] $R_{\text{work}}/R_{\text{free}} = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, where F_{calc} is the calculated protein structure factor from the atomic model (R_{free} was calculated with 5% of the reflections selected).

using PDBsum (Laskowski et al., 2018), which calculated a buried surface area of 1213 Å^2. This is supported by two salt bridges formed between R57 within α4 of protomer A and E110 within α6 of protomer B, and vice versa (Fig. 1H). Hydrogen bonds were calculated by PDBsum as forming between F78 of protomer A and K4 of protomer B (and again vice versa). The rest of the interface is proposed to form through van der Waals interactions. This solved structure shows a stable Aca2 dimer forming a single globular unit with the HTH domains positioned to recognise DNA sequences.

3.2. Analysis of the Aca2 dimer

Next, we examined the surface properties of the Aca2 dimer based on both electrostatic potential (Fig. 2A), and residue conservation (Fig. 2B). Whilst the upper CTD surface of the dimer contains mixed patches of both electropositive and electronegative potential, the NTD HTH motif is clearly electropositive and primed for DNA binding (Fig. 2A, left). When rotated upwards 90° to visualize the "underside" of the Aca2 dimer, there is a clear groove of electropositivity across the entire underside that spans the ~ 30 Å separating the R30 residues and indicates the likely direction for DNA binding (Fig. 2A, right).

ConSurf (Ashkenazy et al., 2016) was used to select sequences homologous to Aca2 (Supplementary Material Table 1), perform a multiple sequence alignment (Supplementary Material Table 2), and then calculate residue conservation from these multiple alignments. This conservation output was then mapped onto the Aca2 surface (Fig. 2B). Interestingly, conservation showed a similar distribution to the electrostatic potential, with greatest conservation in the areas of the HTH and proposed DNA-binding groove, whilst other sections of the α4 backbone helix and CTD were poorly conserved (Fig. 2B).

Previous data have shown that Aca2 autoregulates expression through both DNA binding and bending (Birkholz et al., 2019). Homology modelling via PHYRE 2.0 (Kelley et al., 2015) indicated potential structural homology between DNA-bound MqsA (PDB 3O9X) and Aca2, due to the presence of HTH motifs on both proteins (Birkholz et al., 2019). Having now obtained the Aca2 structure, alignment with MqsA through the HTH domains allowed us to propose a model for Aca2-DNA binding (Fig. 2C). This modelled DNA contains the 20 bp IR1 region of the Aca2 promoter (Fig. 1A) and the two recognition helices can be seen inserting into the major grooves. This also demonstrates how the observed DNA bending might be facilitated by complementary surfaces of Aca2, to allow insertion of the recognition helices (Fig. 2C).

3.3. Structural comparisons of Aca2

The DALI server (Holm and Sander, 1993) was used to search the PDB for structural homologs of Aca2 (Supplementary Material Table 3). The two highest scoring hits were for YdiL from *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT2 (gene *ydiL* aka *STM1362*, PDB 1S4K), and SO3848 from *Shewanella oneidensis* MR-1 (gene *so3848*, PDB 2OX6). YdiL and SO3848 scored Z-scores of 16.6 and 11.9, respectively, and were the only hits that aligned with both domains of Aca2. The bacteria encoding *ydiL* and *so3848* are both γ-proteobacterial pathogens (Heidelberg et al., 2002; McClelland et al., 2001), as is *P. carotovorum*, the host for prophage ZF40 from which Aca2 is derived (Tovkach, 2002). Both the YdiL and the SO3848 structures were produced and deposited by the Midwest Center for Structural Genomics, and both proteins have no known biological function.

Aca2 comprises 116 amino acids, YdiL is 119 amino acids and SO3848 is 166 amino acids. Using EMBOSS Stretcher (Madeira et al., 2019), Aca2 and YdiL share sequence identity of 31.5%, Aca2 and SO3848 share sequence identity of 25.8%, and YdiL and SO3848 share sequence identity of 22.0%, which suggests they are all poorly related to one another at the sequence level. Despite poor sequence similarity, structure-based superposition of Aca2 and YdiL produced an RMSD of 1.8 Å, between 1420 atoms (Fig. 3A). This superposition shows that Aca2 and YdiL are highly similar at the structural level. In contrast, the structure-based superposition of Aca2 and SO3848 is relatively worse, with an RMSD of 3.6 Å, between 1292 atoms (Fig. 3B). Nevertheless, Aca2 and the core regions of SO3848 overlay well (Fig. 3B, left), with variations in secondary structure wherein the equivalent β1 in SO3848 is longer, SO3848 has an additional α-helix between the equivalents of α5 and α6, and β2 is also longer (again forming a parallel β-sheet with β1). SO3848 also has a unique extension to the CTD formed by two additional α-helices that can clearly be seen as additional decorations to the globular core (Fig. 3B, right). These alignments suggest that the biological function of both YdiL and SO3848 is to act as DNA-binding proteins and potential transcriptional regulators, perhaps autorepressors.

We examined the genomic contexts of both YdiL and SO3848 for further clues as to their function (Fig. 3C). Based on genome-wide expression profiling, *ydiL* was not expressed in any of 22 conditions tested in *Salmonella* (Colgan et al., 2016) and is not essential, as determined by TraDIS (Canals et al., 2019). Furthermore, *ydiL* is not part of a prophage (McClelland et al., 2001) or a genomic island, based on an IslandViewer analysis (Bertelli et al., 2017), suggesting the gene is part of the *S.* Typhimurium core genome. Assuming that YdiL is a regulatory protein, it is therefore possible that it binds to sites at distant genomic locations, especially considering its apparent stand-alone character (Fig. 3C) – a stark contrast to the *acrIF8–aca2* operon found in *P. carotovorum* phage ZF40, or *aca* genes in general. However, we also identified several inverted repeats in the vicinity of *ydiL*, including an IR overlapping with the −10 site that might mediate autorepression (Fig. 3C, red inset).

S. oneidensis SO3848 is encoded downstream of a gene encoding a

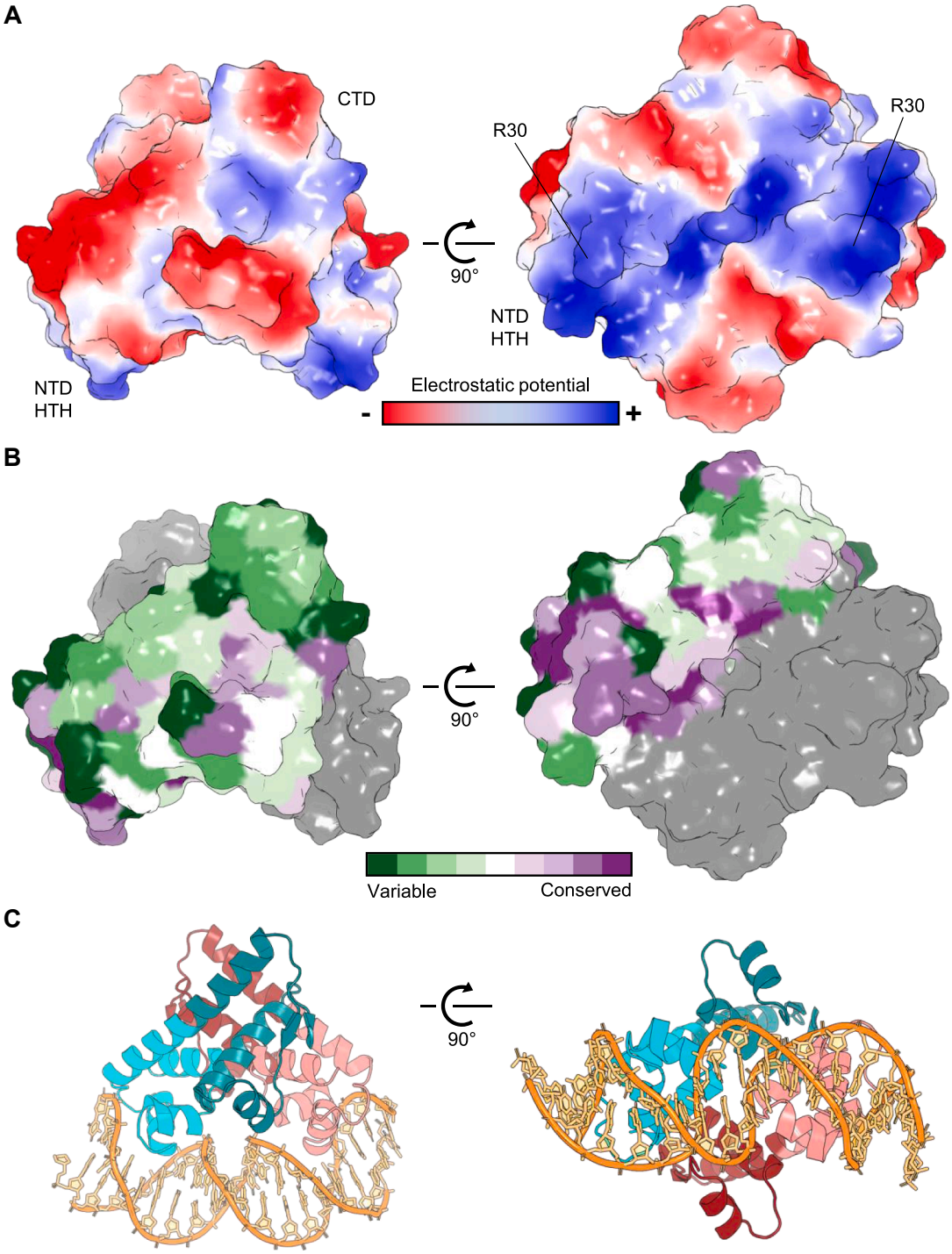


Fig. 2. Analysis of Aca2 dimers. (A) Electrostatic surface potential shows electropositivity (blue) in the NTD HTH domains of Aca2 (left panel). There is an electropositive groove between the two HTH domains and respective key DNA-binding residues from each protomer, R30, are positioned ~ 30 Å apart (right panel). (B) Conservation plots on one Aca2 protomer (colored green to purple as per scale), shown in dimer form (second protomer in gray). (C) Aca2 dimer modelled in complex with 20 bp IR1 dsDNA.

predicted SMI1/KNR4-family protein (Fig. 3C), upstream of which we identified an IR overlapping the −35 site (Fig. 3C, orange inset). Bacterial homologs of the SMI1/KNR4 family have been implicated in contact-dependent inhibition systems (Zhang et al., 2011). Given the context of an Aca2-like regulator, it is possible that this gene has evolved to fulfil an alternative function as an anti-CRISPR; however, similar to *ydiL*, *so3848* does not appear to be part of a genomic island or prophage (as determined using IslandViewer (Bertelli et al., 2017), PHASTER (Arndt et al., 2016) and Prophage Hunter (Song et al., 2019)). From sequence analysis alone it is unclear whether *so3848* and its upstream gene form an operon; despite their close proximity, *so3848* appears to have its own promoter with BPROM-predicted −10 and −35 sites (Solovyev and Salamov, 2011) (Fig. 3C, purple inset). Global profiling showed that *so3848* is expressed in *S. oneidensis* (Kolker et al., 2005) and its expression level is affected by different terminal electron acceptors (Beliaev et al., 2005). The disparate genomic settings of the three Aca2

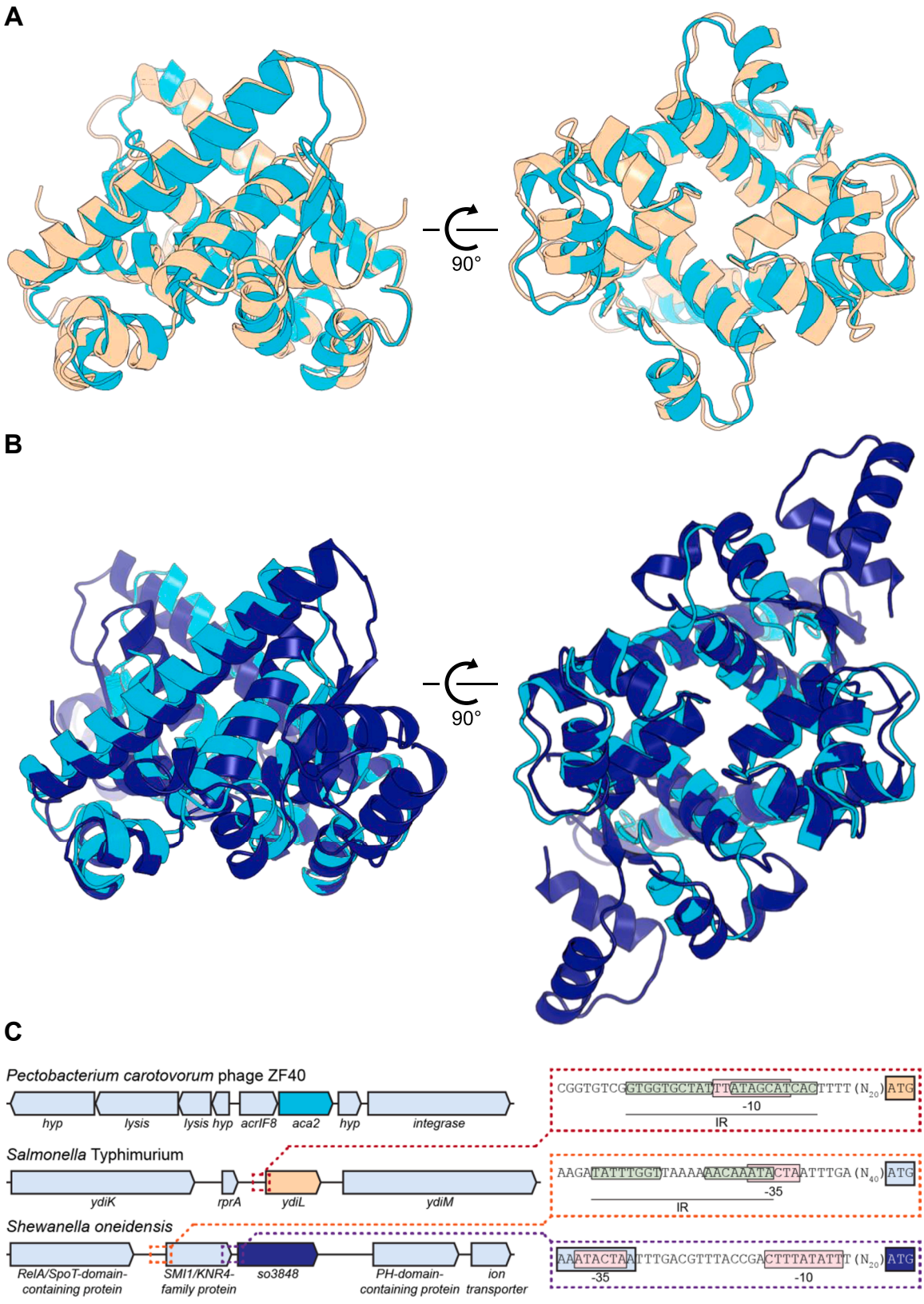


Fig. 3. Structural homologs of Aca2. (A) Superposition of Aca2 dimer (cyan) with YdiL dimer (pale orange, PDB 1S4K). (B) Superposition of Aca2 dimer (cyan) with SO3848 dimer (dark blue, PDB 2OX6). (C) Genomic context of Aca2 structural homologs. *hyp*: gene encoding hypothetical protein. Promoter regions of interest are enlarged on the right; pink boxes indicate promoter elements (-10 and −35 regions), green boxes indicate inverted repeats; coding regions are indicated in the same color as in the overview on the left.

structural homologs (Fig. 3C) suggest diverse implementation of a conserved DNA-binding strategy.

4. Discussion

In this study we have determined the crystal structure of the Aca2 anti-CRISPR-associated transcriptional autorepressor. The obtained structure supports the earlier biological data on DNA binding and bending, with conserved residues lining the electropositive DNA-binding surfaces. We previously proposed that upon phage infection, anti-CRISPR expression initiates strongly and Aca2 will switch off anti-CRISPR production once the host defence has been shut down, potentially to reduce toxic side-effects of AcrIF8 (Birkholz et al., 2019). Should the phage enter lysogeny, Aca2 autoregulation would ensure AcrIF8 levels are suppressed, enabling the host to maintain some CRISPR–Cas activity and ensure protection from secondary infections.

Whilst the Aca2 NTD HTH is highly conserved, the CTD used for dimerization has not been previously characterized, suggesting Aca2 represents a new family of transcriptional regulators. Indeed, Aca2 stands out among the known Aca family members in terms of its larger size (Bondy-Denomy et al., 2018), with other, smaller Aca proteins likely consisting of only one domain (including the HTH motif) and therefore using other dimerization mechanisms. Two structural homologs of Aca2, both from bacterial pathogens, were identified in databases, although both were uncharacterized outputs from structural genomics efforts. Together, the three homologs appear in various genomic contexts, suggesting that the Aca2 family might be more extensively widespread, and that this structural scaffold may be involved in regulating a large range of biological processes. Further work will be needed to fully examine a DNA-bound structure and investigate more diverse members of this nascent family.

5. Accession number

The crystal structure of Aca2 has been deposited in the Protein Data Bank under accession number 7B5J.

CRediT authorship contribution statement

Ben Usher: Investigation, Visualization, Writing - original draft. **Nils Birkholz:** Investigation, Visualization, Writing - original draft. **Izaak N. Beck:** Investigation, Visualization, Writing - original draft. **Robert D. Fagerlund:** Supervision, Writing - review & editing. **Simon A. Jackson:** Supervision, Writing - review & editing. **Peter C. Fineran:** Conceptualization, Funding acquisition, Supervision, Writing - original draft. **Tim R. Blower:** Conceptualization, Funding acquisition, Supervision, Investigation, Visualization, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge Diamond Light Source for time on beamline I04 under proposal MX24948. This work was supported by a Springboard Award from the Academy of Medical Sciences (SBF002\1104) [B.U.], a BBSRC NLD Doctoral Training Partnership studentship (BB/M011186/1) [I.N.B.], a University of Otago Doctoral Scholarship and a Publishing Bursary [N.B.], a University of Otago Research Grant [P.C.F. and R.D.F.], and the Bio-protection Research Centre (Tertiary Education Commission, NZ) [P.C.F.].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2021.107752.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66 (2), 213–221.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44 (W1), W16–W21.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., Ben-Tal, N., 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44 (W1), W344–W350.

Aslanidis, C., de Jong, P.J., 1990. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18 (20), 6069–6074.

Beliaev, A.S., Klingeman, D.M., Klappenbach, J.A., Wu, L., Romine, M.F., Tiedje, J.M., Nealson, K.H., Fredrickson, J.K., Zhou, J., 2005. Global transcriptome analysis of *Shewanella oneidensis* MR-1 exposed to different terminal electron acceptors. J. Bacteriol. 187 (20), 7138–7145.

Bertelli, C., Laird, M.R., Williams, K.P., Lau, B.Y., Hoad, G., Winsor, G.L., Brinkman, F.S., 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 45, W30–W35.

Birkholz, N., Fagerlund, R.D., Smith, L.M., Jackson, S.A., Fineran, P.C., 2019. The autoregulator Aca2 mediates anti-CRISPR repression. Nucleic Acids Res. 47, 9658–9665.

Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., Davidson, A.R., 2013. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature 493 (7432), 429–432.

Bondy-Denomy, J., Davidson, A.R., Doudna, J.A., Fineran, P.C., Maxwell, K.L., Moineau, S., Peng, X.u., Sontheimer, E.J., Wiedenheft, B., 2018. A Unified resource for tracking anti-CRISPR names. CRISPR J. 1 (5), 304–305.

Cai, Y., Usher, B., Gutierrez, C., Tolcan, A., Mansour, M., Fineran, P.C., Condon, C., Neyrolles, O., Genevaux, P., Blower, T.R., 2020. A nucleotidyltransferase toxin inhibits growth of *Mycobacterium tuberculosis* through inactivation of tRNA acceptor stems. Sci. Adv. 6 (31), eabb6651. https://doi.org/10.1126/sciadv.abb6651.

Canals, R., Chaudhuri, R.R., Steiner, R.E., Owen, S.V., Quinones-Olvera, N., Gordon, M. A., Baym, M., Ibba, M., Hinton, J.C.D., Baumler, A.J., 2019. The fitness landscape of the African *Salmonella* Typhimurium ST313 strain D23580 reveals unique properties of the pBT1 plasmid. PLoS Pathog. 15 (9), e1007948. https://doi.org/10.1371/journal.ppat.1007948.

Colgan, A.M., Kröger, C., Diard, M., Hardt, W.-D., Puente, J.L., Sivasankaran, S.K., Hokamp, K., Hinton, J.C.D., Casadesús, J., 2016. The impact of 18 ancestral and horizontally-acquired regulatory proteins upon the transcriptome and sRNA landscape of *Salmonella enterica* serovar typhimurium. PLOS Genet. 12 (8), e1006258. https://doi.org/10.1371/journal.pgen.1006258.

Cowtan, K., 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. Sect. D Biol. Crystallogr. 62 (9), 1002–1011.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. 60 (12), 2126–2132.

Gore, S., Velankar, S., Kleywegt, G.J., 2012. Implementing an X-ray validation pipeline for the Protein Data Bank. Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (4), 478–483.

Hampton, H.G., Watson, B.N.J., Fineran, P.C., 2020. The arms race between bacteria and their phage foes. Nature 577 (7790), 327–336.

He, F., Bhoobalan-Chitty, Y., Van, L.B., Kjeldsen, A.L., Dedola, M., Makarova, K.S., Koonin, E.V., Brodersen, D.E., Peng, X.u., 2018. Anti-CRISPR proteins encoded by archaeal lytic viruses inhibit subtype I-D immunity. Nat. Microbiol. 3 (4), 461–469.

Heidelberg, J.F., Paulsen, I.T., Nelson, K.E., Gaidos, E.J., Nelson, W.C., Read, T.D., Eisen, J.A., Seshadri, R., Ward, N., Methe, B., Clayton, R.A., Meyer, T., Tsapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Haft, D.H., Kolonay, J.F., Madupu, R., Peterson, J.D., Umayam, L.A., White, O., Wolf, A.M., Vamathevan, J., Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J., Khouri, H., Gill, J., Utterback, T.R., McDonald, L.A., Feldblyum, T.V., Smith, H.O., Venter, J.C., Nealson, K.H., Fraser, C.M., 2002. Genome sequence of the dissimilatory metal ion–reducing bacterium *Shewanella oneidensis*. Nat. Biotechnol. 20 (11), 1118–1123.

Holm, L., Sander, C., 1993. Protein structure comparison by alignment of distance matrices. J. Mol. Biol. 233 (1), 123–138.

Kabsch, W., 2010. XDS. Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2), 125–132.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10 (6), 845–858.

Kolker, E., Picone, A.F., Galperin, M.Y., Romine, M.F., Higdon, R., Makarova, K.S., Kolker, N., Anderson, G.A., Qiu, X., Auberry, K.J., Babnigg, G., Beliaev, A.S., Edlefsen, P., Elias, D.A., Gorby, Y.A., Holzman, T., Klappenbach, J.A., Konstantinidis, K.T., Land, M.L., Lipton, M.S., McCue, L.-A., Monroe, M., Pasa-Tolic, L., Pinchuk, G., Purvine, S., Serres, M.H., Tsapin, S., Zakrajsek, B.A., Zhu, W., Zhou, J., Larimer, F.W., Lawrence, C.E., Riley, M., Collart, F.R., Yates, J.R., Smith, R.

D., Giometti, C.S., Nealson, K.H., Fredrickson, J.K., Tiedje, J.M., 2005. Global profiling of *Shewanella oneidensis* MR-1: expression of hypothetical genes and improved functional annotations. Proc. Natl. Acad. Sci. U. S. A. 102 (6), 2099–2104.

Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., Thornton, J.M., 2018. PDBsum: structural summaries of PDB entries. Protein Sci. 27 (1), 129–134.

León, L.M., Park, A.E., Borges, A.L., Zhang, J.Y., Bondy-Denomy, J., 2021. Mobile element warfare via CRISPR and anti-CRISPR in *Pseudomonas aeruginosa*. Nucleic Acids Res. 49, 2114–2125.

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641.

Malone, L.M., Birkholz, N., Fineran, P.C., 2021. Conquering CRISPR: how phages overcome bacterial adaptive immunity. Curr. Opin. Biotech. 68, 30–36.

Marino, N.D., Zhang, J.Y., Borges, A.L., Sousa, A.A., Leon, L.M., Rauch, B.J., Walton, R. T., Berry, J.D., Joung, J.K., Kleinstiver, B.P., Bondy-Denomy, J., 2018. Discovery of widespread type I and type V CRISPR-Cas inhibitors. Science 362 (6411), 240–242.

Marino, N.D., Pinilla-Redondo, R., Csörgő, B., Bondy-Denomy, J., 2020. Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. Nat. Methods 17 (5), 471–479.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., Wilson, R.K., 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature 413 (6858), 852–856.

Osuna, B.A., Karambelkar, S., Mahendra, C., Sarbach, A., Johnson, M.C., Kilcher, S., Bondy-Denomy, J., 2020. Critical anti-CRISPR locus repression by a bi-functional Cas9 inhibitor. Cell Host Microbe 28 (1), 23–30.e5.

Pawluk, A., Staals, R.H.J., Taylor, C., Watson, B.N.J., Saha, S., Fineran, P.C., Maxwell, K. L., Davidson, A.R., Brouns, S.J., Barrangou, R., et al., 2016a. Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. Nat. Microbiol. 321, 16085.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., Davidson, A.R., 2016b. Naturally occurring off-switches for CRISPR-Cas9. Cell 167 (7), 1829–1838.e9.

Pinilla-Redondo, R., Shehreen, S., Marino, N.D., Fagerlund, R.D., Brown, C.M., Sørensen, S.J., Fineran, P.C., Bondy-Denomy, J., 2020. Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. Nat. Commun. 11, 5652.

Rodríguez, D.D., Grosse, C., Himmel, S., González, C., de Ilarduya, I.M., Becker, S., Sheldrick, G.M., Usón, I., 2009. Crystallographic *ab initio* protein structure solution below atomic resolution. Nat. Methods 6 (9), 651–653.

Solovyev, V., Salamov, A., 2011. Automatic annotation of microbial genomes and metagenomic sequences. In: Li, R.W. (Ed.), Metagenomics and Its Applications in Agriculture, Biomedicine and Environmental Studies. (Nova Science Publishers), pp. 62–78.

Song, W., Sun, H.-X., Zhang, C., Cheng, L., Peng, Y., Deng, Z., Wang, D., Wang, Y., Hu, M., Liu, W., et al. (2019). Prophage Hunter: an integrative hunting tool for active prophages. Nucleic Acids Res. 47, W74–W80.

Stanley, S.Y., Borges, A.L., Chen, K.-H., Swaney, D.L., Krogan, N.J., Bondy-Denomy, J., Davidson, A.R., 2019. Anti-CRISPR-associated proteins are crucial repressors of anti-CRISPR transcription. Cell 178 (6), 1452–1464.e13.

Tovkach, F.I., 2002. A study of *Erwinia carotovora* phage resistance with the use of temperate bacteriophage ZF40. Microbiology 71, 72–78.

Wiegand, T., Karambelkar, S., Bondy-Denomy, J., Wiedenheft, B., 2020. Structures and strategies of anti-CRISPR-mediated immune suppression. Annu. Rev. Microbiol. 74 (1), 21–37.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R. M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. 67 (4), 235–242.

Yin, Y., Yang, B., Entwistle, S., 2019. Bioinformatics identification of anti-CRISPR Loci by Using Homology, Guilt-by-Association, and CRISPR Self-Targeting Spacer Approaches. MSystems 4.

Zhang, D., Iyer, L.M., Aravind, L., 2011. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. Nucleic Acids Res. 39, 4532–4552.